cogitatio

POLITICS AND GOVERNANCE

Volume 10

Issue 4

2022

Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics

Open Access Journal

ISSN: 2183-2463

Edited by Alessandro Nai, Diego Garzia, Loes Aaldering, Frederico Ferreira da Silva and Katjana Gattermann

Politics and Governance, 2022, Volume 10, Issue 4
Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Academic Editors
Alessandro Nai (University of Amsterdam)
Diego Garzia (University of Lausanne)
Loes Aaldering (Free University Amsterdam)
Frederico Ferreira da Silva (University of Lausane)
Katjana Gattermann (University of Amsterdam)

Available online at: www.cogitatiopress.com/politicsandgovernance

This issue is licensed under a Creative Commons Attribution 4.0 International License (CC BY). Articles may be reproduced provided that credit is given to the original and *Politics and Governance* is acknowledged as the original venue of publication.

Table of Contents

For a Research Agenda on Negative Politics
Alessandro Nai, Diego Garzia, Loes Aaldering, Frederico Ferreira da Silva, and Katjana Gattermann 243–246

How Partisanship Matters: A Panel Study on the Democratic Outcomes of Perceived Dirty Campaigning
Franz Reiter and Jörg Matthes 247–260

It's All Relative: Perceptions of (Comparative) Candidate Incivility and Candidate Sympathy in Three Multiparty Elections
Chiara Vargiu 261–274

The Decision to Go Negative: Election Types, Candidate Characteristics, and Electoral Competition
Huang-Ting Yan 275–285

The Role of Gender in Parliamentary Attacks and Incivility
Željko Poljak 286–298

The Personality Origins of Positive and Negative Partisanship
Alexa Bankert 299–310

Is Protest Only Negative? Examining the Effect of Emotions and Affective Polarization on Protest Behaviour
Luca Bettarelli, Caroline Close, and Emilie van Haute 311–324

Negative Party Identification and the Use of Party Cues in the Direct Democratic Context
Maxime Walder and Oliver Strijbis 325–335

Fueling Toxicity? Studying Deceitful Opinion Leaders and Behavioral Changes of Their Followers
Puck Guldemond, Andreu Casas Salleras, and Mariken van der Velden 336–348

Who Are the "Dark" Politicians? Insights From Self-Reports of German State Parliament Candidates
Jürgen Maier, Mona Dian, and Corinna Oschatz 349–360

Do Leader Evaluations (De)Mobilize Voter Turnout? Lessons From Presidential Elections in the United States
Liran Harsgor and Neil Nevitte 361–373

Table of Contents

Loud and Negative: Exploring Negativity in Voter Thoughts About Women and Men Politicians
Tobias Rohrbach 374–383

Grievance Politics: An Empirical Analysis of Anger Through the Emotional Mechanism of *Ressentiment*
Tereza Capelos, Mikko Salmela, and Gabija Krisciunaite 384–395

Online Trolls: Unaffectionate Psychopaths or Just Lonely Outcasts and Angry Partisans?
Monika Verbalyte, Christoph Keitel, and Krista Howard 396–410

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 243–246
https://doi.org/10.17645/pag.v10i4.6622

Editorial

For a Research Agenda on Negative Politics

Alessandro Nai [1,*], Diego Garzia [2], Loes Aaldering [3], Frederico Ferreira da Silva [2], and Katjana Gattermann [1]

[1] Amsterdam School of Communication Research, University of Amsterdam, The Netherlands
[2] Institute of Political Studies, University of Lausanne, Switzerland
[3] Department of Political Science and Public Administration, Vrije Universiteit Amsterdam, The Netherlands

* Corresponding author (a.nai@uva.nl)

Submitted: 16 December 2022 | Published: 30 December 2022

Abstract
This thematic issue deals with the "negative" side of politics, more specifically with dynamics of political aggressiveness and ideological opposition in voters and elites. Why do candidates "go negative" on their rivals? To what extent are voters entrenched into opposing camps parted by political tribalism? And are these dynamics related to the (dark) personality of candidates and the expression of emotions in voters? A series of contributions written by leading and emerging scholars provide novel and groundbreaking empirical evidence along three main lines: (a) the evolution, causes, and consequences of political attacks and incivility by political elites; (b) the drivers and dimensions of affective polarization and negative voting in the public; and (c) the dynamics of candidate's personality and perceptions, the affective roots of attitudes and behaviors. This thematic issue aims at setting the stage for a new research agenda on negative politics, able to generate new insights by triangulating evidence and approaches from strands of literature that have mostly evolved on separate tracks.

Keywords
anger; affective polarization; dark personality; incivility; negative campaigning; negative partisanship; negative politics; negative voting; protest; rage; trolling

Issue
This editorial is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This editorial is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

In sharp contrast with the normative ideal of working together toward finding consensual solutions for the greater good, contemporary politics—at least in Western democracies—seems to be built predominantly on oppositional and conflictual elements. At the level of political elites, much attention has been granted in recent years to the dynamics of negative campaigning and attack politics (Nai, 2020). Dealing with an electorate with waning interest in politics, parties and candidates face a strong incentive to "go negative"; as a result, voters are exposed to campaign messages that often include more attacks towards the program and character of the opponent than concrete policy propositions—which might foster cynicism in the electorate. On top of this shift toward attacking the opponent, contemporary politics also seems *qualitatively* more aggressive. Recent research has thoroughly documented the rise of political incivility (Rossini et al., 2021) and the general tendency of elites toward breaking social norms. Trump, Bolsonaro, Duterte, and many more, easily come to mind in this sense. Even more broadly, a clear aggressive stance seems central in the populist worldview, where a normative struggle between the pure people and the evil elite often takes center stage (Hameleers et al., 2018). All in all, political elites seem increasingly drawn toward showcasing negative, confrontational, aggressive behavior—likely due to the rise of political figures with darker personality profiles and characters (Nai & Martínez i Coma, 2019).

But negative politics is not an exclusivity of political elites. In voters as well, politics is often a matter of opposing what they dislike more than striving to promote what they might want or like. Within the electoral arena, growing evidence points towards dynamics of negative voting (Garzia & Ferreira da Silva, 2022; Medeiros & Noël, 2014), where voters cast their choice not to promote candidates or parties they support, but rather to stop candidates or parties they dislike. Even more profoundly, opposition between competing ideas—which, normatively, remains a cornerstone of a healthy democracy—is increasingly supplanted by a "principle dislike" against political foes. Such affective polarization (Iyengar et al., 2019), strongly rooted in dynamics of group identity and tribalism (Mason, 2018), increasingly sets the stage for politics as a war between opposing camps holding irreconcilable moral positions. Today, partisan differences in voters seem almost necessarily to morph into affective polarization and profound dislike of the other camp, which potentially can provide a fertile ground for the development of political violence (Kalmoe & Mason, 2022). Outside of the political arena, political activism often takes the form of contentious mobilization, and waves of mass protests regularly shook the established democratic order.

All in all, contemporary politics seems to be built on an intrinsically negative component. Elites privilege an aggressive stance against each other, echoed by the entrenchment of profound dislike between opposed campaigns in voters—and likely fueled by the obsession with the negative side of politics by news media (Geer, 2012). Yet, surprisingly, research on these phenomena has mostly evolved on separate tracks. This thematic issue takes stock of these separate strands of research and brings together empirical work on election campaigning, leader personality, negative voting, and antagonistic political attitudes towards the goal to start setting the stage for an integrated framework on Negative Politics.

2. Contributions

The contributions in our thematic issue can be classified into three broad topics: (a) drivers and consequences of negativity in election campaigns, (b) the roots of affective polarization and negative voting, and (c) the dynamics of candidate personality, perceptions, and emotions.

Starting with election campaigns, Reiter and Matthes (2022) introduce the concept of "dirty campaigns"—that is, campaigns that violate social norms by, e.g., engaging in incivility and deceitful campaigning techniques. On top of expanding our conceptual toolkit when it comes to thinking about how political elites engage in "negative" campaigns, their article goes a step further and shows how such "dirty" campaigns can lead to distrustful attitudes in the public, likely moderated by partisan attitudes. Vargiu (2022) offers a novel take on political incivility and argues for the necessity to account for voters' perceptions of such incivility. Following a constructionist perspective, the author looks at how such perceptions shape candidate likeability during recent elections in France, Germany, and the Netherlands—and shows that perceived incivility tends to correspond to more negative feelings towards candidates, but also that it is relative incivility, more than absolute levels, that does the heavy lifting when it comes to candidate sympathy. Yan (2022) looks at recent elections in Taiwan (2008–2022) to uncover drivers of negative campaigning at the candidate and competition levels and highlights the importance of contextual factors when it comes to modeling the decision to go negative. Beyond expanding our understanding of the drivers of negativity in such an important and overlooked case, the article relies on a methodological approach—qualitative comparative analysis—rarely used in communication research. Poljak (2022) investigates the presence of attacks and incivility during "routine times," looking at parliamentary debates in Belgium, Croatia, and the UK. The author focuses specifically on gender dynamics, and shows evidence that politicians tend to adhere to gender stereotypes—women attack less (and are less likely to use incivility) and are more rarely targeted by attacks.

Turning to affective polarization and negative voting, using a sample of American and Swedish respondents, Bankert (2022) investigates the influence of personality traits (e.g., the "Big Five," authoritarianism, etc.) on negative and positive partisanship. Results show strongly heterogeneous effects, indicating that the personality origins of partisanship differ across countries (and party affiliation)—suggesting the need for more comparative research. Bettarelli et al. (2022) bridge the gap between the literature on emotions, affective polarization, and protest behavior. Looking at survey data from Belgian voters, the authors uncover the affective roots of political perceptions and actions, for instance, by showing how anger and hope towards politics can effectively combine to drive voters towards engaging in protest behavior, and how affective polarization can compensate for the lack of such emotions. Walder and Strijbis (2022) look at the use of party cues within the context of Swiss direct democratic votes, focusing in particular on the effects of negative party identification. Triangulating experimental and observational evidence, the authors show that voters tend indeed to align themselves against parties they dislike, which opens up an important new line of research on negative partisanship during referenda. Guldemond et al. (2022) investigate the extent to which Dutch opinion leaders on Twitter spread deceiving content and the effects that such content has. Via the computational analysis of a large sample of tweets, the authors show that users who "follow" a deceitful opinion leader become more affectively polarized.

Finally, turning to dynamics of candidate perceptions, personality, and emotions, Maier et al. (2022) offer one of the very first insights into the "dark" personality of politicians that relies on self-ratings from candidates running for German state elections—opening up

an exciting new research agenda that tackles the nefarious nature of politicians' character via what candidates themselves say about their own profile. Harsgor and Nevitte (2022) investigate whether evaluations of presidential candidates drive turnout in American elections. Using long-term survey data (1968–2020), the authors show that turnout is affected by the differences in affect between the main competing candidates, and the nature of such affect. Rohrbach (2022) dives into how negativity is expressed in voters' thoughts about women and men politicians when exposed to negative media cues—and how these thoughts affect, in turn, their vote preferences. Results across two studies with German-speaking respondents suggest that negative cues generate negativity in voters' thoughts similarly for men and women, but such negative thoughts seem more prejudicial for the electoral chances of men. Capelos et al. (2022) dive into the psychology of the "angry voter." A close look at interviews with "angry" American citizens reveals the centrality of *ressentiment*—that is, the tendency to transform grievances (e.g., injustice, shame, envy) into anti-social outputs associated with morally righteous indignation, rage, and destructive anger. Verbalyte et al. (2022) provide a sociological explanation of "trolls" who engage in personal attacks and insults online. Looking at a sample of American respondents, the authors show the existence of two main categories of trolling: one based on fun and entertainment and another with more defensive/reactive roots. Personal profiles, such as political identity and religiosity, play a role in whether people engage in such trolling activities online.

Acknowledgments

We are very grateful to the authors for their excellent—and groundbreaking—contributions, and to the pool of reviewers for their critical and constructive insights that guided us throughout the process of establishing this thematic issue. Several of the articles in the thematic issue were discussed during the first workshop on negative politics that we organized online in late 2021—thank you to all participants for their precious input.

Conflict of Interests

The authors declare no conflict of interests.

References

Bankert, A. (2022). The personality origins of positive and negative partisanship. *Politics and Governance*, *10*(4), 299–310.

Bettarelli, L., Close, C., & van Haute, E. (2022). Is protest only negative? Examining the effect of emotions and affective polarization on protest behaviour. *Politics and Governance*, *10*(4), 311–324.

Capelos, T., Salmela, M., & Krisciunaite, G. (2022). Grievance politics: An empirical analysis of anger through the emotional mechanism of *ressentiment*. *Politics and Governance*, *10*(4), 384–395.

Garzia, D., & Ferreira da Silva, F. (2022). The electoral consequences of affective polarization? Negative voting in the 2020 US Presidential election. *American Politics Research*, *50*(3), 303–311.

Geer, J. G. (2012). The news media and the rise of negativity in presidential campaigns. *PS: Political Science & Politics*, *45*(3), 422–427.

Guldemond, P., Salleras, A. C., & van der Velden, M. (2022). Fueling toxicity? Studying deceitful opinion leaders and behavioral changes of their followers. *Politics and Governance*, *10*(4), 336–348.

Hameleers, M., Bos, L., & de Vreese, C. (2018). Framing blame: Toward a better understanding of the effects of populist communication on populist party preferences. *Journal of Elections, Public Opinion and Parties*, *28*(3), 380–398.

Harsgor, L., & Nevitte, N. (2022). Do leader evaluations (de)mobilize voter turnout? Lessons from presidential elections in the United States. *Politics and Governance*, *10*(4), 361–373.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhortra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146.

Kalmoe, N. P., & Mason, L. (2022). *Radical American partisanship: Mapping violent hostility, its causes, and the consequences for democracy.* University of Chicago Press.

Maier, J., Dian, M., & Oschatz, C. (2022). Who are the "dark" politicians? Insights from self-reports of German state parliament candidates. *Politics and Governance*, *10*(4), 349–360.

Mason, L. (2018). *Uncivil agreement: How politics became our identity.* University of Chicago Press.

Medeiros, M., & Noël, A. (2014). The forgotten side of partisanship: Negative party identification in four Anglo-American democracies. *Comparative Political Studies*, *47*(7), 1022–1046.

Nai, A. (2020). Going negative, worldwide: Towards a general understanding of determinants and targets of negative campaigning. *Government and Opposition*, *55*(3), 430–455.

Nai, A., & Martínez i Coma, F. (2019). The personality of populists: Provocateurs, charismatic leaders, or drunken dinner guests? *West European Politics*, *42*(7), 1337–1367.

Poljak, Ž. (2022). The role of gender in parliamentary attacks and incivility. *Politics and Governance*, *10*(4), 286–298.

Reiter, F., & Matthes, J. (2022). How partisanship matters: A panel study on the democratic outcomes of perceived dirty campaigning. *Politics and Governance*, *10*(4), 247–260.

Rohrbach, T. (2022). Loud and negative: Exploring negativity in voter thoughts about women and men politicians. *Politics and Governance*, *10*(4), 374–383.

Rossini, P., Stromer-Galley, J., & Zhang, F. (2021). Exploring the relationship between campaign discourse on Facebook and the public's comments: A case study of incivility during the 2016 US presidential election. *Political Studies*, *69*(1), 89–107.

Vargiu, C. (2022). It's all relative: Perceptions of (comparative) candidate incivility and candidate sympathy in three multiparty elections. *Politics and Governance*, *10*(4), 261–274.

Verbalyte, M., Keitel, C., & Howard, C. (2022). Online trolls: Unaffectionate psychopaths or just lonely outcasts and angry partisans? *Politics and Governance*, *10*(4), 396–410.

Walder, M., & Strijbis, O. (2022). Negative party identification and the use of party cues in the direct democratic context. *Politics and Governance*, *10*(4), 325–335.

Yan, H.-T. (2022). The decision to go negative: Election types, candidate characteristics, and electoral competition. *Politics and Governance*, *10*(4), 275–285.

About the Authors

Alessandro Nai is an assistant professor of political communication at the Amsterdam School of Communication Research (ASCoR), University of Amsterdam. His recent work deals with the dark sides of politics, the use of negativity and incivility in election campaigns from a comparative perspective, and the (dark) personality traits of political figures. He is currently directing a research project that maps the use of negative campaigning in elections across the world. His most recent book is *Dark Politics: The Personality of Politicians and the Future of Democracy* (Oxford University Press, 2023, with J. Maier). He is currently associate editor of the *Journal of Elections, Public Opinion & Parties* and the *Journal of Social and Political Psychology*.

Diego Garzia is an assistant professor of political science at the University of Lausanne. His academic interests center on the interplay between voters, political parties, and leaders in democratic elections. His current research project focuses on affective polarization and its electoral consequences, with a particular emphasis on negative voting from comparative and longitudinal perspectives. He currently serves on the Steering Committees of the Italian National Election Study (ITANES) and the Consortium of National Election Studies (CNES).

Loes Aaldering is an assistant professor in comparative politics at the Department of Political Science and Public Administration of the Vrije Universiteit Amsterdam. She is a board member of the Dutch Political Science Association (NKWP) and a member of the steering committee of the ECPR Standing Group "Political Communication." Her research interests include gender and politics, political leadership, political representation, and electoral behavior. In her current projects, she studies (gendered) mediatized leadership effects on voters, gender differences in media coverage of politicians, political gender stereotypes, fake news, news avoidance, and polarization.

Frederico Ferreira da Silva has a PhD from the European University Institute (2019). He is a senior SNSF researcher at the University of Lausanne, Switzerland. He is also a research fellow in the European Governance and Politics Programme of the Robert Schuman Centre for Advanced Studies at the European University Institute. His research focuses on elections, public opinion, voting behavior, and voting advice applications.

Katjana Gattermann is an assistant professor of political communication and journalism at the Amsterdam School of Communication Research (ASCoR) and is also affiliated with the Amsterdam Centre for European Studies (ACES). She is member of the steering committee of the ECPR Standing Group "Political Communication." Her research interests comprise political communication, political behaviour, public opinion, journalism, and legislative behaviour, with a regional focus on the European Union. Her most recent book is *The Personalization of Politics in the European Union* (Oxford University Press, 2022).

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 247–260
https://doi.org/10.17645/pag.v10i4.5672

Article

How Partisanship Matters: A Panel Study on the Democratic Outcomes of Perceived Dirty Campaigning

Franz Reiter * and Jörg Matthes

Department of Communication, University of Vienna, Austria

* Corresponding author (franz.reiter@univie.ac.at)

Submitted: 27 April 2022 | Accepted: 20 September 2022 | Published: 30 December 2022

Abstract
Uncivil campaigning and deceitful campaign techniques are increasingly relevant phenomena in politics. However, it remains unclear how they share an underlying component and how partisanship can influence their associations with democratic outcomes. We introduce the concept of dirty campaigning, which is situated at the intersection of research on negative campaigning and political scandals. Dirty campaigning involves violations of social norms and liberal-democratic values between elite political actors in terms of style and practices, such as uncivil campaigning and deceitful campaign techniques. In a two-wave panel study (N = 634) during the 2021 German federal election campaign, we investigate the associations of perceived dirty campaigning by the least and most favorite party with distrust in politicians, trust in democracy, attitudes toward dirty campaigning regulation, as well as perceived harmful consequences of dirty campaigning for democracy. We find that perceived dirty campaigning by the least favorite party increases perceptions of harmful consequences of dirty campaigning for democracy over time. In contrast, perceived dirty campaigning by the most favorite party decreases perceptions of harmful consequences of dirty campaigning for democracy as well as attitudes toward dirty campaigning regulation over time. Perceptions of harmful consequences of dirty campaigning for democracy increase distrust in politicians over time and vice versa. Our findings suggest that the outcomes of dirty campaigning can depend on partisanship and can have important implications for the quality of democracy.

Keywords
democratic outcomes; dirty campaigning; panel study; political incivility; political trust

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Uncivil campaigning and deceitful campaign techniques are prevalent phenomena in politics. Extant research has shown that they can contribute to a toxic political environment, undermine the integrity of elections, and harm the quality of democracy (Geer, 2006; Mattes & Redlawsk, 2014; Mutz, 2015; Sydnor, 2019; Walter, 2021). Despite the progress made by previous research, two research gaps remain. First, it is unclear how uncivil campaigning and deceitful campaign techniques share an underlying component. Prior work has advanced the conceptualization of these phenomena (e.g., Brooks & Geer, 2007; Hinds et al., 2020; Lösche, 2019; Stryker et al., 2016), but conceived them as independent of each other rather than developing a complementary framework to investigate them as a coherent concept. Second, there is a lack of research on how partisanship can influence the democratic outcomes of uncivil campaigning and deceitful campaign techniques. Previous studies on their democratic outcomes ignored the role of partisanship and instead focused on personal dispositions,

different communication channels, or different countries (Mutz, 2015; Otto et al., 2020; Walter & Vliegenthart, 2010). However, studying partisanship in this context is important, because it can lead to variability in individual reactions to campaigns (Druckman et al., 2019; Fridkin & Kenney, 2011). Research on the role of partisanship investigated negative campaigning as an umbrella term rather than specifically uncivil campaigning and deceitful campaign techniques (e.g., Haselmayer et al., 2020; Nai, 2013; Somer-Topcu & Weitzel, 2022).

To address these research gaps, we introduce the concept of dirty campaigning as actions between elite political actors that violate social norms and values of liberal democracy in terms of style and practices and may include uncivil campaigning as well as deceitful campaign techniques (Hinds et al., 2020; Mutz, 2015; Stryker et al., 2021). In a two-wave panel survey (*N* = 634), we examine how perceived dirty campaigning by the least and most favorite party is associated with distrust in politicians, trust in democracy, attitudes toward dirty campaigning regulation, and perceived harmful consequences of dirty campaigning for democracy.

2. A Basic Conceptualization of Dirty Campaigning

Uncivil campaigning refers to the use of incivility between political elites (Chen, 2017; Mutz, 2015) and is conceived to be a subform of negative campaigning (Brooks & Geer, 2007; Haselmayer, 2019). Uncivil campaigning may involve forms of insults (i.e., name-calling or disrespect), discourse (i.e., interrupting political opponents), modality (i.e., sarcasm or ambiguity), or context (i.e., calls for political violence; Bormann et al., 2021; Mutz, 2015; Stryker et al., 2016, 2021). In this work, we rely on insult utterances and discursive forms as the most widely shared conceptualizations of uncivil campaigning (Mutz, 2015; Stryker et al., 2016, 2021). Therefore, we regard uncivil campaigning as communication of an elite political actor A against an elite political actor B, which includes norm violations in terms of utterances and discursive forms.

Deceitful campaign techniques involve non-communicative practices that are unethical and disproportionate. Research on such techniques is rather fragmented and broadly embedded in the literature on political scandals. Forms of deceitful campaign techniques may involve illegal campaign financing (Lösche, 2019), financing of news media for favorable coverage (Dragomir, 2017), or the creation of deepfakes (Meskys et al., 2020). We label this form of dirty campaigning deceitful campaign "techniques" instead of deceitful campaign "methods." Campaign techniques imply a technical action, whereas campaign methods could also relate to the systematic use of dirty forms of campaigning, such as uncivil campaigning. We thus refer to deceitful campaign techniques as the use of non-communicative practices by an elite political actor A against an elite political actor B, which are unethical or disproportionate.

We argue that these forms share the underlying notion of campaign forms that violate social norms and values of liberal democracy in terms of style and practices. Uncivil campaigning violates the social norms of civil style in interpersonal exchanges (Mutz, 2015; Stryker et al., 2016; Walter, 2021), whereas deceitful campaign techniques violate the social norms of using practices that are ethical or proportionate (Gächter & Schulz, 2016). These norm violations can involve different degrees of severity and have different outcomes. Nevertheless, in the first place, they all have in common that they involve norm violations.

This argument can be extended to violations of the values of liberal democracy. Uncivil campaigning is considered to undermine democratic civility by contributing to a toxic political atmosphere and impairing public discourse (Chen, 2017; Flores et al., 2021). Deceitful campaign techniques are regarded to violate the principle of political integrity because they constitute an abuse of political power to achieve unethical or disproportionate goals (Grant, 1999; Thompson, 2000). Previous research demonstrated that uncivil campaigning and deceitful campaign techniques tend to have more negative than positive implications for democratic quality (Walter, 2021). Uncivil campaigning can decrease trust in politicians, congress, and government, as well as political participation intentions and policy support (Fridkin & Kenney, 2019; Mutz, 2015; Otto et al., 2020; Reiter & Matthes, 2021; Van 't Riet & Van Stekelenburg, 2021). Deceitful campaign techniques can decrease positive evaluations and voting intentions toward the sponsor, as well as decrease trust in politicians and trust in democracy (Vivyan et al., 2012; Von Sikorski et al., 2020). Taken together, we define dirty campaigning as actions of an elite political actor A against an elite political actor B that violate social norms and values of liberal democracy in terms of style and practices.

Our concept of dirty campaigning provides a complementary framework, which is located at the intersection of research on negative campaigning and political scandals. For instance, uncivil campaigning may be conceived both as a subform of negative and dirty campaigning. It may be investigated within the theoretical framework of negative campaigning (i.e., civil and uncivil campaigning) and dirty campaigning (i.e., uncivil campaigning and deceitful campaign techniques). Deceitful campaign techniques may constitute a subform of political scandals rather than negative campaigning because they involve non-communicative practices instead of communicative actions between political actors.

2.1. Dirty Campaigning in Germany

Previous research has shown that dirty forms of campaigning are a permanent component of federal campaigns in Germany (Hopmann et al., 2018; Maier & Renner, 2018). However, compared to other countries, the amount of dirty campaigning appears to be relatively

low and has declined over time (Schmücking, 2015; Walter, 2014). The reasons for this development can be rooted in the multi-party system, which decreases the likelihood to use dirty campaigning against potential coalition partners (Elmelund-Præstekær, 2010), as well as the strongly consensus-oriented political culture in Germany, which fosters democratic civility (Lijphart, 1999).

Despite these developments, the rise of the Alternative of Germany (AfD), a right-wing populist party, and their entry into the German Bundestag in 2017 have raised concerns that dirty campaigning could become increasingly relevant and that the electorate may become more polarized (Arzheimer & Berning, 2019; Nai, 2018). Evidence from previous research showed that speeches of the AfD in the German Bundestag contained more uncivil campaigning than speeches of the other parliamentary factions. In turn, the share of uncivil campaigning in speeches of the other parliamentary factions increased (Maurer & Jost, 2020).

Dirty campaigning was also common in the 2021 German federal election campaign (Dostal, 2021). For instance, a private company ran a false poster campaign against the Greens, associating them with "eco-terror" or "climate socialism" (Ruppert, 2021). The Social Democratic Party (SPD) used uncivil campaigning against the frontrunner of the Christian Democratic Union (CDU), Armin Laschet. They addressed one of his party members' conservative religious beliefs, thus breaking a taboo in German campaigns (Monath, 2021). A survey also showed that the majority of German citizens perceived the 2021 campaign to be too aggressive (Gensing, 2021). This lends some evidence that dirty campaigning is still a prevalent phenomenon in German campaigns.

3. Dirty Campaigning and Democratic Outcomes

We investigate four democratic outcomes associated with dirty campaigning: distrust in politicians, trust in democracy, attitudes toward dirty campaigning regulation, as well as perceived harmful consequences of dirty campaigning on the quality of democracy. Distrust in politicians is a very specific form of political trust, which involves the lack of confidence in politicians' ability to "do what is right" (Miller & Listhaug, 1990, p. 358), to be unresponsive to public needs, or to be unethical (Easton, 1975; Warren, 2017). Trust in democracy is a diffuse form of political trust, which comprises support for democratic principles and values, as well as evaluations of the performance of democracies (Norris, 2011). Both forms are important indicators of democratic quality because citizens need to have faith in the policymaking of their elected representatives and the effective functioning of democracy (Hetherington, 2004; Miller & Listhaug, 1990). Attitudes toward dirty campaigning regulation address the individual evaluation that the use of dirty campaigning requires stronger action from legislators, such as the creation or amendment of laws, or the strengthening of previous laws. This aspect is democratically relevant, because the regulation of harmful campaign behavior, such as dirty campaigning, is an important legal instrument of a democracy (Marsden et al., 2020). The consequences of dirty campaigning for democratic quality are important to study from the perception of citizens (Lipsitz & Geer, 2017) and involve the perceived consequences for the integrity of elections and effective policymaking (Graham & Svolik, 2020; Norris, 2011).

3.1. The Outcomes of Perceived Dirty Campaigning by the Least and Most Favorite Party

The new videomalaise theory (NVT; Mutz, 2015) argues that citizens perceive the use of uncivil campaigning as a violation of social norms, which can negatively influence attitudes toward politicians and democratic processes. Regarding the outcomes of perceived dirty campaigning by the least favorite party, the social identity theory (SIT; see Hogg, 2016; Tajfel & Turner, 1986) suggests that individuals assess their social identities by comparing their ingroup to specific outgroups. When individuals consider that their in-group status is made salient by the actions of an out-group, group categorizations are activated. Consequently, individuals may use heuristics (i.e., undesirable actions of an out-group) to develop evaluations toward the out-group (Branscombe & Wann, 1994). Partisanship may constitute such an important social identity that can affect how individuals evaluate actions by a political party they consider as least favorite (Druckman et al., 2013; Iyengar & Westwood, 2015). A further explanation offers the directional goal of motivated reasoning theory. According to this theory, individuals are motivated to reach desired conclusions by giving greater weight to attitude-consistent information than attitude-challenging information (Kunda, 1990). Partisan-motivated reasoning can occur when individuals are primed to draw conclusions that are consistent with their party identification (Taber & Lodge, 2006). In other words, individuals may tend to support and favor actions by their most favorite party, whereas they oppose and dislike actions by their least favorite party.

Based on these theories, individuals may regard perceived dirty campaigning by the least favorite party as an undesirable out-group action. This may negatively influence democratic outcomes and increase distrust in politicians as well as decrease trust in democracy (Hetherington, 2004; Norris, 2011). Perceived dirty campaigning by the least favorite party may arguably increase citizens' desire for stronger regulation of dirty campaigning. Individuals may also perceive dirty campaigning by the least favorite party to be harmful to electoral integrity and effective policymaking, which may increase the perceived harmful consequences of dirty campaigning for democratic quality (Mutz, 2015; Norris, 2014; Taber & Lodge, 2006). We thus hypothesize:

H1: Perceived dirty campaigning by the least favorite party (a) increases distrust in politicians and (b) decreases trust in democracy over time.

H2: Perceived dirty campaigning by the least favorite party increases (a) attitudes toward dirty campaigning regulation and (b) perceptions of harmful consequences of dirty campaigning for democracy over time.

The NVT would suggest that perceived dirty campaigning by the most favorite party is considered a breach of social norms (Mutz, 2015), whereas SIT would assume that certain actions by this party are more accepted than those by other parties (Muddiman & Stroud, 2017; Tajfel & Turner, 1986). In this context, previous research showed inconclusive findings. For instance, uncivil campaigning by the most favorite party is not related to attitudes toward this party (Gervais, 2019), whereas uncivil partisan media can increase negative attitudes toward the most favorite party (Druckman et al., 2019). Given this inconclusive evidence and the conflicting assumptions of the NVT and SIT, the associations of variables of interest could arguably produce null findings, be less negative compared to the least favorite party, or even be positive. We thus pose a research question:

RQ1: How is perceived dirty campaigning by the most favorite party associated with distrust in politicians and trust in democracy over time?

RQ2: How is perceived dirty campaigning by the most favorite party associated with attitudes toward dirty campaigning regulation and perceptions of harmful consequences of dirty campaigning for democracy over time?

3.2. The Outcomes of Perceived Harmful Consequences of Dirty Campaigning for Democracy

Based on input-performance approaches of democratic theory (Hakhverdian & Mayne, 2015; Scharpf, 1999), citizens who perceive dirty campaigning to be harmful to democratic quality might associate dirty campaigning with politicians as their sponsors and thus have decreasing levels of confidence in them (Norris, 2014). Citizens with harmful perceptions of dirty campaigning for democratic quality may arguably have little faith in the performance of a democracy to effectively counter dirty campaigning, and thus lose trust in democracy (Norris, 2011). Furthermore, they could prefer a stronger regulation of dirty campaigning by lawmakers to limit its harmful democratic consequences (Meskys et al., 2020). We thus postulate:

H3: Perceptions of harmful consequences of dirty campaigning for democracy increase (a) distrust in politicians, (b) decrease trust in democracy, and (c) increase attitudes toward dirty campaigning regulation over time.

3.3. The Outcomes of Political Trust

According to the trust-as-heuristic thesis (Rudolph, 2017), citizens who distrust politicians do not believe them “to do what is right” (Miller & Listhaug, 1990, p. 358) and thus may advocate that dirty campaigning by politicians requires stronger regulation. They may also perceive that their use of dirty campaigning impairs democratic performance and thus perceive dirty campaigning to have negative consequences for democratic quality (Norris, 2014; Warren, 2017). Citizens with high levels of trust in democracy may assume that democracies are sufficiently responsive to regulate dirty campaigning, which decreases their individual need for further regulation (Marsden et al., 2020; Norris, 2011). Similarly, they may believe that effective democratic regime performance would diminish the perceived negative consequences of dirty campaigning for the quality of democracy (Hetherington, 2004). We postulate:

H4: Distrust in politicians increases (a) attitudes toward dirty campaigning regulation and (b) perceptions of harmful consequences of dirty campaigning for democracy over time.

H5: Trust in democracy decreases (a) attitudes toward dirty campaigning regulation and (b) perceptions of harmful consequences of dirty campaigning for democracy over time.

Figure 1 depicts our hypothesized model.

4. Method

4.1. Procedure

We conducted a two-wave online panel survey during the 2021 German federal election campaign. Ethical approval for this study was obtained from the Institutional Review Board of the Department of Communication, University of Vienna (approval ID: 20210713_053). Data are openly available (Reiter & Matthes, 2022). Dynata, a professional polling company, collected the survey data at two points in time between 29 July–4 August 2021 and 13–22 September 2021, with the election date on 26 September 2021.

4.2. Sample

We used a quota sample of the German electoral population based on age (ranging from 18 to 80 years, $M = 53.41$, $SD = 14.27$), gender (48.4% of the respondents identified as female), and education (13.7% lower education, 56.6% medium education, 29.7% higher education). To ensure high data quality, we excluded “speeders”

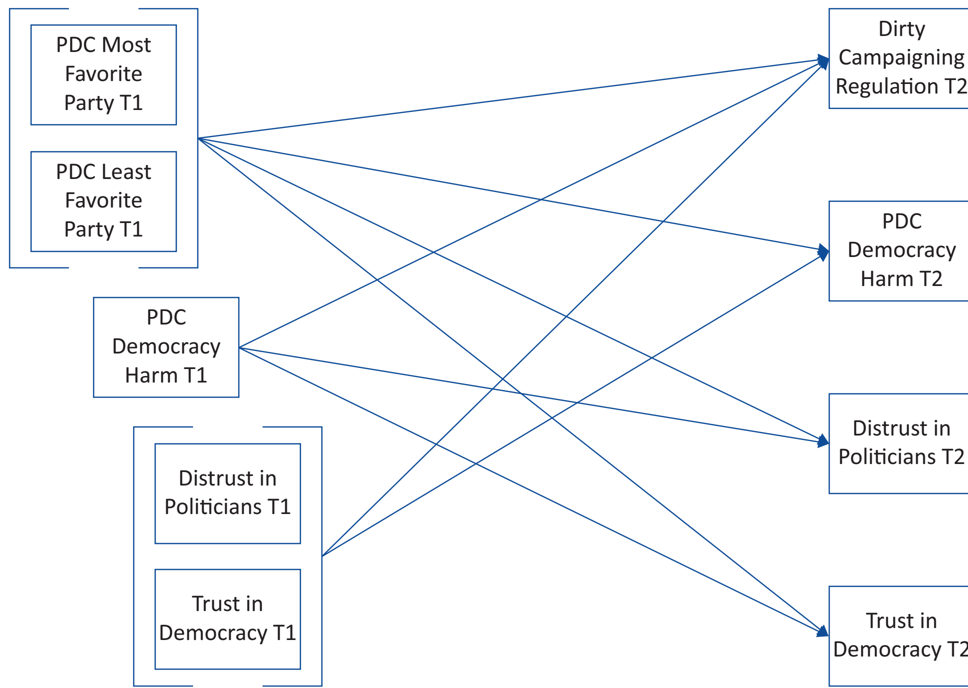

Figure 1. Hypothesized model. Notes: PDC stands for "perceived dirty campaigning"; square brackets indicate arrows of each independent variable on dependent variables.

and performed attention checks (for complete details of excluded responses see the Supplementary File, Table A1). Our final sample size was *N* = 634. The retention rate of the responses of the final samples between wave one and wave two was 67.66% (for complete details of systematic differences of samples between both waves, see Supplementary File, Table A2).

4.3. Measures

Frequency distribution of the least and most favorable party are reported in Table 1. Complete details of the descriptive statistics for our measures are reported in Table 2. We employ McDonald's Omega for reliability estimation of three or more items. We use the OMEGA macro for SPSS with Hancock's algorithm (HA) and 1,000 bootstrapping samples to generate 95% confidence intervals (CI; Hayes & Coutts, 2020). For reliability estimation of two items, we use the Spearman-Brown coefficient. If not stated differently, we employed a seven-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*) for the measurements of the variables.

To measure *perceived dirty campaigning by the least favorite party*, we created a filter variable asking the participants about their least favorite party in the German Bundestag, followed by two items to assess their perception of dirty campaigning by the selected party (items based on Reiter & Matthes, 2021): "The [FILTER PARTY] is disrespectful to other parties"; "The [FILTER PARTY] uses deceitful campaign techniques, for instance illegal campaign financing." We then computed a new variable consisting of that mean value per participant. For *perceived dirty campaigning by the most favorite party*, we applied the same procedure by asking about the most

Table 1. Frequency distribution of the least and most favorite party.

	Least favorite party		Most favorite party	
Political party	T1	T2	T1	T2
CDU/CSU	34	29	170	131
SPD	12	12	116	160
AfD	410	397	85	94
FDP	14	15	83	88
The Left	41	60	74	77
The Greens	123	121	106	84
Total	634	634	634	634

Note: T1 stands for "Time 1" and T2 for "Time 2."

favorite party. We measured *attitudes toward dirty campaigning regulation* with three items (based on Nelson et al., 2021): "Dirty campaigning should legally be more regulated"; "The sponsors of dirty campaigning should be prosecuted more vigorously"; "Dirty campaigning should be legally penalized more strongly." To measure *perceptions of harmful consequences of dirty campaigning for democracy*, we used six items (derived from Norris, 2014), two of which were: "Dirty campaigning makes election campaigns look manipulated"; "Dirty campaigning contributes to a hostile political atmosphere." We measured *distrust in politicians* with four items (derived from Craig et al., 1990), two of which were: "Politicians in Germany are more concerned with their own interests than with actual policies"; "Politicians in Germany rarely keep their promises to the people." We measured *trust in democracy* with four items (based on Norris, 2011), two of which were: "Democracy is the right form of government for Germany"; "I am satisfied with the way democracy works in Germany." To close any potential "back-door paths" which may influence the association between the variables of interest, we purposefully controlled for *demographics* (age, gender, educational level), *ideology,* and *perceived civil negative campaigning* (Rohrer, 2018; for complete details see Supplementary File, Table A3).

4.4. Data Analysis

We ran four OLS regression models (model 1: $R^2_{Adj.}$ = .34, $F(12, 621) = 27.99$, $p < .001$; model 2: $R^2_{Adj.}$ = .42, $F(12, 621) = 39.24$, $p < .001$; model 3: $R^2_{Adj.}$ = .56, $F(12, 621) = 68.77$, $p < .001$; model 4: $R^2_{Adj.}$ = .62, $F(12, 621) = 88.07$, $p < .001$) with lagged dependent variables. We also included autoregressive paths to explain changes in the dependent variable from T1 to T2. We observed no model specification error (Ramsey, 1969) in all four models (model 1: $F(3, 618) = 1.15$, $p = .327$; model 2: $F(3, 618) = 2.51$, $p = .058$; model 3: $F(3, 618) = 1.05$, $p = .372$; model 4: $F(3, 618) = 0.17$, $p = .919$). We also detected no indication of multicollinearity, as the VIF-values for the predictors in all four models were reported lower than 2.5 (Alin, 2010).

Table 2. Descriptive statistics of measures.

Variable	Reliability coefficient	Mean	Standard deviation
Perceived dirty campaigning least favorite party	T1: ρ = .81 T2: ρ = .77	T1: M = 5.15 T2: M = 4.90	T1: SD = 1.69 T2: SD = 1.66
Perceived dirty campaigning most favorite party	T1: ρ = .73 T2: ρ = .69	T1: M = 2.51 T2: M = 2.43	T1: SD = 1.31 T2: SD = 1.32
Attitudes toward dirty campaigning regulation	T1: ω = .93, 95% CI [.92, .95] T2: ω = .94, 95% CI [.92, .95]	T1: M = 5.43 T2: M = 5.41	T1: SD = 1.42 T2: SD = 1.44
Perceptions of harmful consequences of dirty campaigning for democracy	T1: ω = .91, 95% CI [.89, .92] T2: ω = .90, 95% CI [.88, .92]	T1: M = 5.31 T2: M = 5.36	T1: SD = 1.28 T2: SD = 1.26
Distrust in politicians	T1: ω = .88, 95% CI [.86, .90] T2: ω = .89, 95% CI [.87, .91]	T1: M = 4.86 T2: M = 4.87	T1: SD = 1.45 T2: SD = 1.47
Trust in democracy	T1: ω = .75, 95% CI [.69, .79] T2: ω = .72, 95% CI [.65, .78]	T1: M = 4.73 T2: M = 4.77	T1: SD = 1.35 T2: SD = 1.36
Age	—	T1: M = 53.41 T2: M = 53.41	T1: SD = 14.27 T2: SD = 14.28
Gender	—	T1: M = 1.48 T2: M = 1.49	T1: SD = .50 T2: SD = .50
Medium education	—	T1: M = .57 T2: M = .58	T1: SD = .50 T2: SD = .49
High education	—	T1: M = .30 T2: M = .31	T1: SD = .46 T2: SD = .46
Ideology	—	T1: M = 4.73	T1: SD = 1.75
Political interest	T1: ρ = .87 T2: ρ = .85	T1: M = 5.61 T2: M = 5.59	T1: SD = 1.49 T2: SD = 1.47
Perceived civil negative campaigning	T1: ω = .87, 95% CI [.85, .89] T2: ω = .88, 95% CI [.86, .90]	T1: M = 3.90 T2: M = 3.98	T1: SD = 1.18 T2: SD = 1.19

5. Results

Results of a t-test revealed a significant difference in the means of perceived dirty campaigning by the least favorite party at T1 (M = 5.15, SD = 1.69) compared to the most favorite party at T1 (M = 2.51, SD = 1.31, $t(633)$ = 32.14, $p < .001$).

5.1. Hypotheses Tests

Table 3 and Figure 2 depict our results. H1a and H1b were rejected because we found no significant association of perceived dirty campaigning by the least favorite party at T1 with distrust in politicians at T2 and trust in democracy at T2. H2a was not supported because the results revealed no significant association of perceived dirty campaigning by the least favorite party at T1 with attitudes toward dirty campaigning regulation at T2. H2b was confirmed by showing a significant positive association of perceived dirty campaigning by the least favorite party at T1 with perceptions of harmful consequences of dirty campaigning for democracy at T2 (see Figure 3). An increase (or decrease) by one *SD* in perceived dirty campaigning by the least favorite party at T1 increases (or decreases) the predicted value for perceptions of harmful consequences of dirty campaigning for democracy at T2 by 0.10 units. Regarding RQ1, the results revealed no significant association of perceived

Table 3. Results of OLS regression analysis.

	Dirty campaigning regulation (T2)	Perceived dirty campaigning democracy harm (T2)	Distrust in politicians (T2)	Trust in democracy (T2)
Female (T1)	0.10 (0.10)	0.12 (0.08)	−0.04 (0.08)	−0.06 (0.07)
Age (T1)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.00 (0.00)
Medium education (T1)	−0.20 (0.14)	0.06 (0.12)	−0.11 (0.12)	0.03 (0.10)
High education (T1)	−0.22 (0.15)	0.04 (0.13)	−0.20 (0.13)	0.00 (0.11)
Ideology (T1)	−0.04 (0.03)	−0.02 (0.02)	0.02 (0.02)	−0.05** (0.02)
Political interest (T1)	0.06 (0.03)	0.01 (0.03)	−0.01 (0.03)	0.10*** (0.02)
Perceived civil negative campaigning (T1)	—	—	−0.08* (0.04)	0.06 (0.03)
Perceived dirty campaigning most favorite party (T1)	−0.08* (0.04)	−0.10*** (0.03)	0.06 (0.03)	0.00 (0.03)
Perceived dirty campaigning least favorite party (T1)	0.02 (0.03)	0.06** (0.02)	0.02 (0.02)	0.01 (0.02)
Dirty campaigning regulation (T1)	0.44*** (0.04)	0.13*** (0.04)	—	—
Perceived dirty campaigning democracy harm (T1)	0.12* (0.05)	0.44*** (0.04)	0.09* (0.03)	0.03 (0.03)
Distrust in Politicians (T1)	0.11** (0.04)	0.09** (0.03)	0.64*** (0.03)	−0.05 (0.03)
Trust in democracy (T1)	0.04 (0.04)	0.03 (0.04)	−0.14*** (0.04)	0.69*** (0.03)
Constant	1.35** (0.48)	1.20** (0.39)	1.86*** (0.42)	0.93** (0.36)
Adj. R^2	0.34	0.42	0.56	0.62
F	27.99	39.24	68.77	88.07
N	634	634	634	634

Notes: Standard errors in parentheses; * $p < .05$, ** $p < .01$, *** $p < .001$.

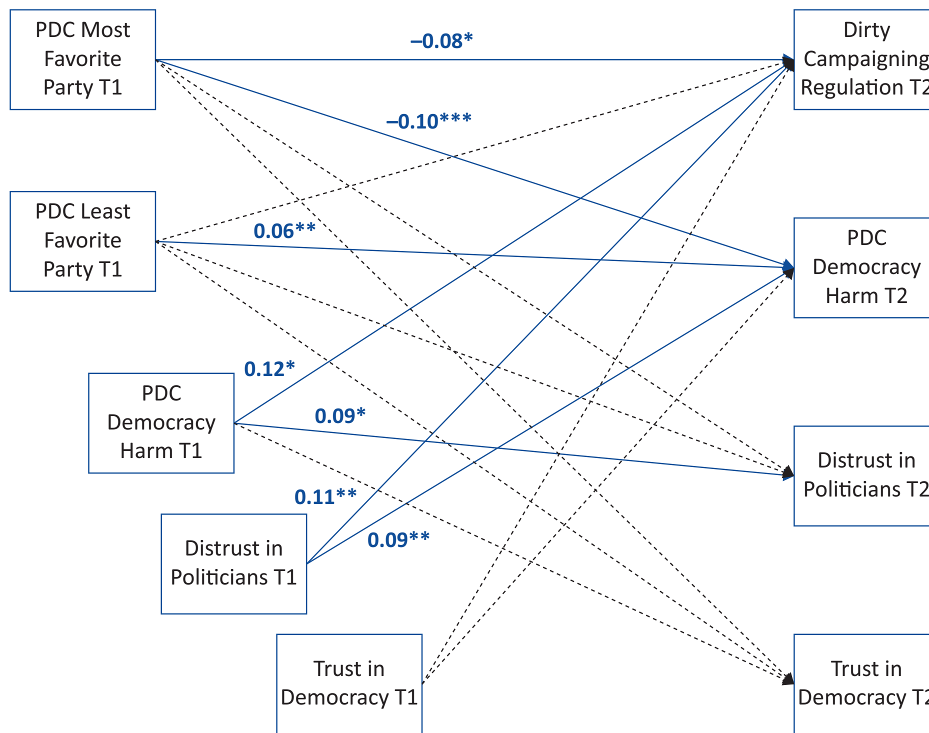

Figure 2. Associations of hypothesized model. Notes: Figure based on Table 3; * $p < .05$, ** $p < .01$, *** $p < .001$.

dirty campaigning by the most favorite party at T1 with distrust in politicians at T2 and trust in democracy at T2.

For RQ2, the results indicated a significant negative association of perceived dirty campaigning by the most favorite party at T1 with attitudes toward dirty campaigning regulation at T2 and perceptions of harmful consequences of dirty campaigning for democracy at T2 (see Figure 3). An increase (or decrease) by one *SD* in perceived dirty campaigning by the most favorite party at T1 decreases (or increases) the predicted value for

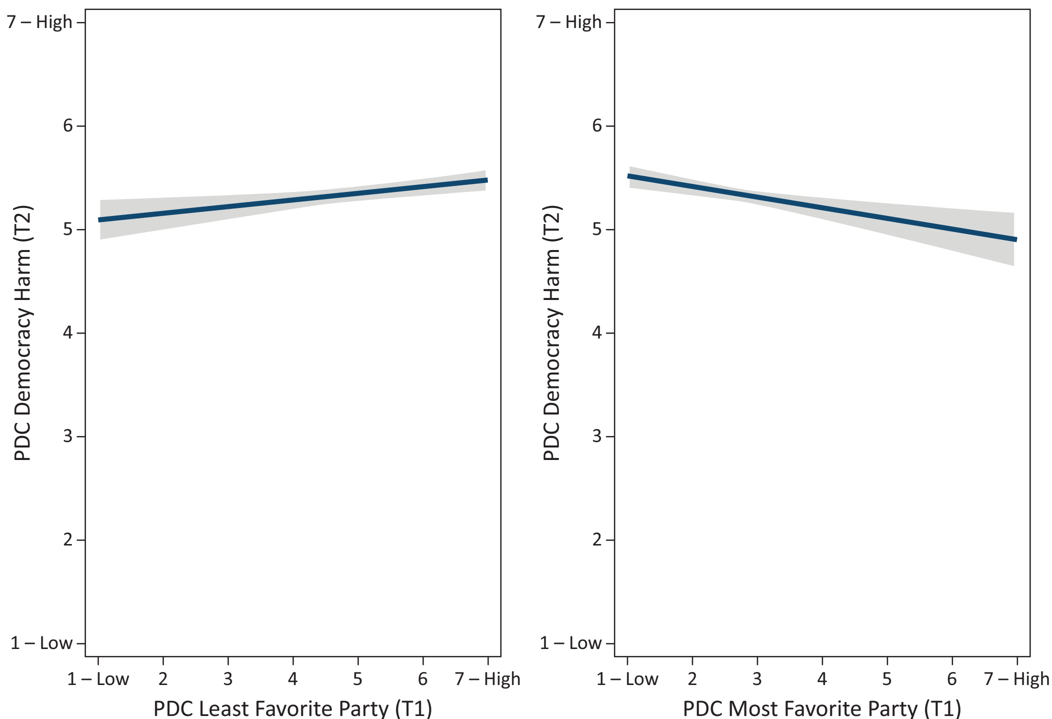

Figure 3. Associations of perceived dirty campaigning by the least favorite party (left graph) and the most favorite party (right graph) with perceptions of harmful consequences of dirty campaigning for democracy. Notes: Grey area represents 95% confidence interval; figure based on Table 3.

attitudes toward dirty campaigning regulation at T2 by 0.10 units. An increase (or decrease) by one *SD* in perceived dirty campaigning by the most favorite party at T1 decreases (or increases) the predicted value for perceptions of harmful consequences of dirty campaigning for democracy at T2 by 0.13 units.

Confirming H3a and H3c, we observed a significant positive association of perceptions of harmful consequences of dirty campaigning for democracy at T1 with distrust in politicians at T2 (see Figure 4) and attitudes toward dirty campaigning regulation at T2. H3b was rejected because the results revealed no significant association of perceptions of harmful consequences of dirty campaigning for democracy at T1 with trust in democracy at T2. H4a and H4b were confirmed because we observed a significant positive association of distrust in politicians at T1 with attitudes toward dirty campaigning regulation at T2 and with perceptions of harmful consequences of dirty campaigning for democracy at T2 (see Figure 4).

H5a and H5b were rejected because trust in democracy at T1 showed no significant associations with attitudes toward dirty campaigning regulation at T2 and perceptions of harmful consequences of dirty campaigning for democracy at T2.

6. Discussion

Uncivil campaigning and deceitful campaign techniques are increasingly relevant in politics and have received growing attention in research. However, it is unclear how they share an underlying component and how partisanship may affect their associations with democratic outcomes. In this article, we introduce the concept of dirty campaigning, defined as actions between elite political actors that violate social norms and values of liberal democracy in terms of style and practices (Grant, 1999; Mutz, 2015; Stryker et al., 2016), and may involve uncivil campaigning and deceitful campaign techniques. In a two-wave panel study during the 2021 German federal election campaign, we investigated how perceived dirty campaigning by the least and most favorite party is associated with democratic outcomes. Dirty campaigning in Germany has become more prevalent due to rise of the AfD and their entry into the German Bundestag, thus making it an important case to study (Maurer & Jost, 2020; Walter, 2014).

When discussing findings, we should not only focus on significant p-values but also the size of the beta coefficients (Funder & Ozer, 2019). For our study, these effect sizes involve values from –.08 to .12, which are generally considered to be small (Ferguson, 2009). However, they do not occur at the level of single events but indicate change over time. In the context of our study, these findings demonstrate that perceived dirty campaigning by the least and most favorite can influence democratic outcomes over the course of a campaign. Furthermore, tests for model specification error and multicollinearity prove the robustness of our findings. Our findings may be generalized to other countries to a certain extent, because we investigated individual perceptions of campaigns on the micro level instead of objective characteristics, such as dirty campaigning by political parties, on

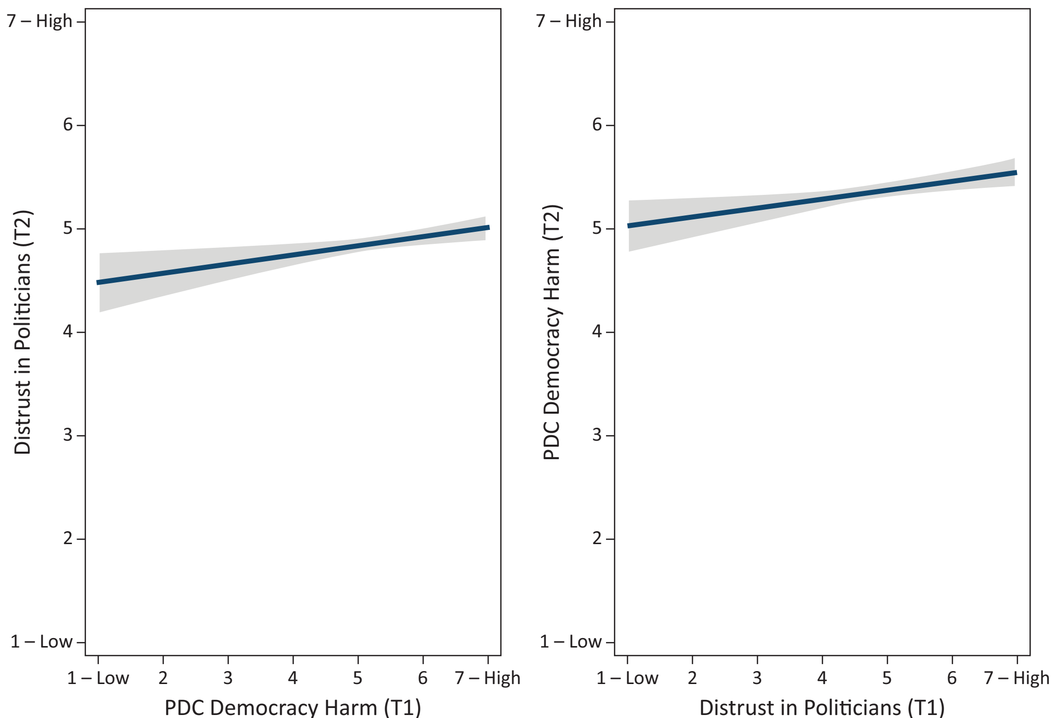

Figure 4. Associations of perceptions of harmful consequences of dirty campaigning for democracy with distrust in politicians (left graph) and vice versa (right graph). Notes: Grey area represents 95% confidence interval; figure based on Table 3.

the mesolevel. Objective characteristics and systemic factors may influence individual perceptions about the campaign, but ultimately these individual perceptions shape the outcomes of dirty campaigning.

Regarding our hypothesized associations, we find that perceived dirty campaigning by the least favorite party increases perceived harmful consequences of dirty campaigning for democratic quality but is not associated with attitudes toward dirty campaigning regulation over time. Perceived dirty campaigning by the most favorite party decreases perceived harmful consequences of dirty campaigning for democracy and attitudes toward dirty campaigning regulation over time. These findings demonstrate that perceived dirty campaigning by the least favorite party may constitute a violation of social norms by an out-group party (Mutz, 2015; Tajfel & Turner, 1986), which is perceived to be harmful to the quality of democracy (Norris, 2014). In contrast, perceived dirty campaigning by the most favorite appears to outweigh the violation of social norms in favor of in-group party thinking (Hogg, 2016). Thus, they may associate dirty campaigning by the most favorite party with a decreasing need for the regulation of dirty campaigning and positive consequences for the quality of democracy.

Perceived dirty campaigning by the least and most favorite party is not directly associated with outcomes related to political trust. We may argue that citizens do not link dirty campaigning to diffuse levels of political support, such as trust in democracy (Hetherington, 2004; Norris, 2011). Instead, they may turn to more specific levels of political support like distrust in politicians. Although we lack the data to investigate mediated associations, we found a positive association of distrust in politicians with perceived harmful consequences of dirty campaigning for democracy. This may suggest that the associations of perceived dirty campaigning by the least and most favorite party with distrust in politicians can be mediated by perceived harmful democratic consequences of dirty campaigning.

Our findings also show that perceived harmful consequences of dirty campaigning for democratic quality increase distrust in politicians and vice versa over time. This suggests, following the input-performance approach (Hakhverdian & Mayne, 2015; Scharpf, 1999), that citizens associate the harmful consequences of dirty campaigning for democratic quality with politicians as sponsors, which increases distrust in them. In turn, distrust in politicians may function as a heuristic for increasing perceptions of harmful democratic consequences of dirty campaigning (Rudolph, 2017). These findings may suggest a reciprocal relationship between perceived harmful consequences of dirty campaigning for democracy and distrust in politicians over time. Furthermore, we find that perceived harmful consequences of dirty campaigning for democratic quality increases attitudes toward dirty campaigning regulation. This finding can also be explained by the input-performance approach, because citizens who perceive dirty campaigning to be harmful to democratic quality may advocate a stronger regulation of dirty campaigning (Marsden et al., 2020). Also, distrust in politicians predicts stronger attitudes toward dirty campaigning regulation. Citizens who distrust politicians may not have the confidence that politicians avoid undesirable behavior, such as dirty campaigning, and therefore demand stronger regulation (Rudolph, 2017). These findings are important because they demonstrate that specific forms of political trust can be associated with attitudes and perceptions toward dirty campaigning and may even form reciprocal associations.

Contrary to our assumptions, perceived harmful consequences of dirty campaigning for democratic quality are not associated with trust in democracy. We may speculate that such perceptions can be related to more specific instead of diffuse levels of political trust (Norris, 2011). Our findings also indicate non-significant associations of trust in democracy with attitudes toward dirty campaigning regulation and the perceived harmful democratic consequences of dirty campaigning. We may speculate that more specific instead of diffuse forms of political trust are associated with regulating politicians as sponsors of dirty campaigning and perceptions of harmful democratic consequences of dirty campaigning (Easton, 1975; Hetherington, 2004).

Our study contributes to previous research by paving the way for a theoretical framework of dirty campaigning. It can be understood as a complementary framework, which is situated at the intersection of research on negative campaigning and political scandals. Our conceptualization still leaves space for further development, as it may go beyond uncivil campaigning and deceitful campaign techniques and involve other facets we have not considered in this study. Nevertheless, our concept provides a first approach under which general aspects of campaign actions among elite political actors may count as dirty.

We also contribute to previous research by demonstrating that the outcomes of dirty campaigning may not be uniform across citizens and that individual variation can depend on partisanship (Druckman et al., 2019; Iyengar & Westwood, 2015). From a micro-level perspective, partisans in their perception may find dirty campaigning by their most favorite party to be beneficial for the quality of a democracy. However, from a macro level perspective—that is the implications for key components of the quality of a democracy, such as electoral integrity and effective policymaking (Lijphart, 1999; Norris, 2014)—these findings appear rather concerning than beneficial. Citizens may downplay dirty campaigning by their most favored party and overrate dirty campaigning by their least favorite party, which can amplify partisan biased information processing (Taber & Lodge, 2006). Citizens may also adopt a partisan "double standard" by forgiving norm-violating behavior and democratic transgressions of their most favorite party compared to other parties (Graham & Svolik, 2020; Simonovits et al., 2022).

6.1. Limitations and Future Research

As with all survey research, we measured individual perceptions instead of actual behavior. This means that respondents may over- or underestimate the extent of dirty campaigning, although we statistically controlled for important third variables and autoregressive associations. However, when studying the outcomes of dirty campaigning, the underlying logic is that only subjective impressions of respondents matter, as they shape how respondents think and act. Also, our study involved two panel waves, which does not allow us to test complex mediation paths across time or to examine within- and between-person effects (Hamaker et al., 2015). Therefore, future research should involve experimental designs or studies with three or more panel waves. Moreover, cross national-research in Western Europe and beyond is highly warranted. The conceptualization of dirty campaigning may also involve facets other than uncivil campaigning or deceitful campaign techniques, which could be investigated in future studies.

7. Conclusion

Dirty campaigning has become increasingly relevant in recent years and there is empirical evidence that such forms can foster democratic backsliding. Our findings from a two-wave panel study demonstrate that partisanship can be important to study the democratically relevant outcomes of dirty campaigning. Citizens tend to perceive dirty campaigning by the least favorite party as harmful and dirty campaigning by the most favorite party as beneficial for the quality of democracy. Although these findings may suggest that dirty campaigning can have positive democratic outcomes in the perception of citizens, this can hold problematic implications for the quality of democracy. Citizens may downgrade dirty campaigning by their most favorite party and overrate dirty campaigning by their least favorite party, thus indicating a partisan perceptual bias. Political parties may use dirty campaigning to make electoral gains, which can contribute to a more toxic political climate.

Acknowledgments

Funding for data collection and open access publication was provided by the University of Vienna.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(3), 370–374. https://doi.org/10.1002/WICS.84

Arzheimer, K., & Berning, C. C. (2019). How the Alternative for Germany (AfD) and their voters veered to the radical right, 2013–2017. *Electoral Studies*, *60*. https://doi.org/10.1016/J.ELECTSTUD.2019.04.004

Bormann, M., Tranow, U., Vowe, G., & Ziegele, M. (2021). Incivility as a violation of communication norms: A typology based on normative expectations toward political communication. *Communication Theory*, *32*(3). https://doi.org/10.1093/CT/QTAB018

Branscombe, N. R., & Wann, D. L. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European Journal of Social Psychology*, *24*(6), 641–657. https://doi.org/10.1002/EJSP.2420240603

Brooks, D. J., & Geer, J. G. (2007). Beyond negativity: The effects of incivility on the electorate. *American Journal of Political Science*, *51*(1), 1–16. https://doi.org/10.1111/j.1540-5907.2007.00233.x

Chen, G. M. (2017). *Online incivility and public debate: Nasty talk*. Palgrave Macmillan.

Craig, S. C., Niemi, R. G., & Silver, G. E. (1990). Political efficacy and trust: A report on the NES pilot study items. *Political Behavior*, *12*(3), 289–314.

Dostal, J. M. (2021). Germany's federal election of 2021: Multi-crisis politics and the consolidation of the six-party system. *The Political Quarterly*, *92*(4), 662–672. https://doi.org/10.1111/1467-923X.13080

Dragomir, M. (2017). Control the money, control the media: How government uses funding to keep media in line: *Journalism*, *19*(8), 1131–1148. https://doi.org/10.1177/1464884917724621

Druckman, J. N., Gubitz, S. R., Levendusky, M. S., & Lloyd, A. M. (2019). How incivility on partisan media (De)polarizes the electorate. *Journal of Politics*, *81*(1), 291–295. https://doi.org/10.1086/699912

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, *107*(1), 57–79. https://doi.org/10.1017/S0003055412000500

Easton, D. (1975). A re-assessment of the concept of political support. *British Journal of Political Science*, *5*(4), 435–457. https://doi.org/10.1017/S0007123400008309

Elmelund-Præstekær, C. (2010). Beyond American negativity: Toward a general understanding of the determinants of negative campaigning. *European Political Science Review*, *2*(1), 137–156. https://doi.org/10.1017/S1755773909990269

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538. https://

doi.org/10.1037/A0015808
Flores, M., Nair, M., Rasmussen, M., & Sydnor, E. (2021). Civility through the comparative lens: Challenges and achievements. In A. S. Walter (Ed.), *Political incivility in the parliamentary, electoral and media arena: Crossing boundaries* (pp. 19–34). Routledge.
Fridkin, K. L., & Kenney, P. (2011). Variability in citizens' reactions to different types of negative campaigns. *American Journal of Political Science*, *55*(2), 307–325. https://doi.org/10.1111/j.1540-5907.2010.00494.x
Fridkin, K. L., & Kenney, P. J. (2019). *Taking aim at attack advertising: Understanding the impact of negative campaigning in U.S. Senate races*. Oxford University Press. https://doi.org/10.1093/oso/9780190947569.001.0001
Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202
Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature 2016*, *531*(7595), 496–499. https://doi.org/10.1038/nature17160
Geer, J. G. (2006). *In defense of negativity: Attack ads in presidential campaigns*. University of Chicago Press.
Gensing, P. (2021). *Desinformation zur Bundestagswahl: Mehrheit befürchtet Beeinflussung* [Disinformation on the bundestag election: Majority fears influence]. Tagesschau. https://www.tagesschau.de/faktenfinder/podcast/desinformation-bundestagswahl-103.html
Gervais, B. T. (2019). Rousing the partisan combatant: Elite incivility, anger, and antideliberative attitudes. *Political Psychology*, *40*(3), 637–655. https://doi.org/10.1111/pops.12532
Graham, M. H., & Svolik, M. W. (2020). Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *American Political Science Review*, *114*(2), 392–409. https://doi.org/10.1017/S0003055420000052
Grant, R. W. (1999). *Hypocrisy and integrity: Machiavelli, Rousseau, and the ethics of politics*. University of Chicago Press.
Hakhverdian, A., & Mayne, Q. (2015). Institutional trust, education, and corruption: A micro-macro interactive approach. *The Journal of Politics*, *74*(3), 739–750. https://doi.org/10.1017/S0022381612000412
Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/A0038889
Haselmayer, M. (2019). Negative campaigning and its consequences: a review and a look ahead. *French Politics*, *17*(3), 355–372. https://doi.org/10.1057/s41253-019-00084-8
Haselmayer, M., Hirsch, L., & Jenny, M. (2020). Love is blind. Partisanship and perception of negative campaign messages in a multiparty system. *Political Research Exchange*, *2*(1). https://doi.org/10.1080/2474736X.2020.1806002
Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than cronbach's alpha for estimating reliability. But... *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629
Hetherington, M. J. (2004). *Why trust matters: Declining political trust and the demise of american liberalism*. Princeton University Press.
Hinds, J., Williams, E. J., & Joinson, A. N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, *143*. https://doi.org/10.1016/J.IJHCS.2020.102498
Hogg, M. A. (2016). Social identity theory. In S. McKeown, R. Haji, & N. Ferguson (Eds.), *Understanding peace and conflict through social identity theory* (pp. 3–17). Springer. https://doi.org/10.1007/978-3-319-29869-6_1
Hopmann, D. N., Vliegenthart, R., & Maier, J. (2018). The effects of tone, focus, and incivility in election debates. *Journal of Elections, Public Opinion and Parties*, *28*(3), 283–306. https://doi.org/10.1080/17457289.2017.1394310
Iyengar, S., & Westwood, S. J. (2015). Fear and Loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707. https://doi.org/10.1111/ajps.12152
Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480
Lijphart, A. (1999). *Patterns of democracy: Government forms and performance in thirty-six countries*. Yale University Press.
Lipsitz, K., & Geer, J. G. (2017). Rethinking the concept of negativity: An empirical approach. *Political Research Quarterly*, *70*(3), 577–589. https://doi.org/10.1177/1065912917706547
Lösche, P. (2019). Problems of party and campaign financing in Germany and the United States—Some comparative reflections. In A. B. Gunlicks (Ed.), *Campaign and party finance in North America and Western Europe* (pp. 219–230). Routledge.
Maier, J., & Renner, A. M. (2018). When a man meets a woman: Comparing the use of negativity of male candidates in single- and mixed-gender televised debates. *Political Communication*, *35*(3), 433–449. https://doi.org/10.1080/10584609.2017.1411998
Marsden, C., Meyer, T., & Brown, I. (2020). Platform values and democratic elections: How can the law regulate digital disinformation? *Computer Law and Security Review*, *36*. https://doi.org/10.1016/j.clsr.2019.105373
Mattes, K., & Redlawsk, D. P. (2014). *The positive case for negative campaigning*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226202334.001.0001

Maurer, M., & Jost, P. (2020). Das Ende der Debattenkultur? Wie die AfD Negativität und Inzivilität in den Bundestag trägt und wie die anderen Parteien darauf reagieren [The end of debate culture? How the AfD brings negativity and incivility into the Bundestag and how the other parties react]. *Zeitschrift Für Politik*, *67*(4), 473–490. https://doi.org/10.5771/0044-3360-2020-4-473

Meskys, E., Kalpokiene, J., Jurcys, P., & Liaudanskas, A. (2020). Regulating deep fakes: Legal and ethical considerations. *Journal of Intellectual Property Law & Practice*, *15*(1), 24–31. https://papers.ssrn.com/abstract=3497144

Miller, A. H., & Listhaug, O. (1990). Political parties and confidence in government: A comparison of Norway, Sweden and the United States. *British Journal of Political Science*, *20*(3), 357–386. https://doi.org/10.1017/S0007123400005883

Monath, H. (2021, August 7). Laschet als Ziel von "Negative Campaigning": Die SPD bricht ein Tabu im Wahlkampf [Laschet as target of "negative campaigning": The SPD breaks a taboo in the election campaign]. *Der Tagesspiegel*. https://www.tagesspiegel.de/politik/laschet-als-ziel-von-negative-campaigning-die-spd-bricht-ein-tabu-im-wahlkampf/27490902.html

Muddiman, A., & Stroud, N. J. (2017). News values, cognitive biases, and partisan incivility in comment sections. *Journal of Communication*, *67*(4), 586–609. https://doi.org/10.1111/JCOM.12312

Mutz, D. C. (2015). *In-your-face politics: The consequences of uncivil media*. Princeton University Press.

Nai, A. (2013). What really matters is which camp goes dirty: Differential effects of negative campaigning on turnout during Swiss federal ballots. *European Journal of Political Research*, *52*(1), 44–70. https://doi.org/10.1111/j.1475-6765.2012.02060.x

Nai, A. (2018). Fear and loathing in populist campaigns? Comparing the communication style of populists and non-populists in elections worldwide. *Journal of Political Marketing*, *20*(2), 219–250. https://doi.org/10.1080/15377857.2018.1491439/SUPPL_FILE/WPLM_A_1491439_SM9785.PDF

Nelson, M. R., Ham, C. D., Haley, E., & Chung, U. C. (2021). How political interest and gender influence persuasion knowledge, political information seeking, and support for regulation of political advertising in social media. *Journal of Interactive Advertising*, *21*(3), 225–242. https://doi.org/10.1080/15252019.2021.1978352

Norris, P. (2011). *Democratic deficit: Critical citizens revisited*. Cambridge University Press.

Norris, P. (2014). *Why electoral integrity matters*. Cambridge University Press.

Otto, L. P., Lecheler, S., & Schuck, A. R. T. (2020). Is context the key? The (non-)differential effects of mediated incivility in three European countries. *Political Communication*, *37*(1), 88–107. https://doi.org/10.1080/10584609.2019.1663324

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, *31*(2), 350–371. https://doi.org/10.1111/j.2517-6161.1969.tb00796.x

Reiter, F., & Matthes, J. (2021). "The good, the bad, and the ugly." A panel study on the reciprocal effects of negative, dirty, and positive campaigning on political distrust. *Mass Communication and Society*. https://doi.org/10.1080/15205436.2021.1934702

Reiter, F., & Matthes, J. (2022). *How partisanship matters: A Panel study on the democratic outcomes of perceived dirty campaigning*. OSF Home. https://osf.io/zpuhr/?view_only%20=%20da82faf765f449d0a220264639e4082f

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rudolph, T. J. (2017). Political trust as a heuristic. In S. Zmerli & T. Van der Meer (Eds.), *Handbook on political trust* (pp. 197–211). Edward Elgar Publishing.

Ruppert, J. (2021). *Anti-Grünen-Kampagne: Wahlkampf nach amerikanischem Vorbild?* [Anti-green campaign: American-style election campaign?]. BR24. https://www.br.de/nachrichten/deutschland-welt/anti-gruenen-kampagne-wahlkampf-nach-amerikanischem-vorbild,Sg9lbIY

Scharpf, F. (1999). *Governing in Europe: Effective and democratic?* Oxford University Press.

Schmücking, D. (2015). *Negative Campaigning: Die Wirkung und Entwicklung negativer politischer Werbung in der Bundesrepublik* [Negative campaigning: The impact and development of negative political advertising in the Federal Republic of Germany]. Springer. https://doi.org/10.1007/978-3-658-08212-3

Simonovits, G., McCoy, J., & Littvay, L. (2022). Democratic hypocrisy and out-group threat: Explaining citizen support for democratic erosion. *The Journal of Politics*, *84*(3). https://doi.org/10.1086/719009

Somer-Topcu, Z., & Weitzel, D. (2022). Negative campaigning and vote choice in europe: How do different partisan groups react to campaign attacks? *Comparative Political Studies*. Advance online publication. https://doi.org/10.1177/00104140221074283

Stryker, R., Conway, B. A., Bauldry, S., & Kaul, V. (2021). Replication note: What is political incivility? *Human Communication Research*, *48*(1), 168–177. https://doi.org/10.1093/hcr/hqab017

Stryker, R., Conway, B. A., & Danielson, J. T. (2016). What is political incivility? *Communication Monographs*, *83*(4), 535–556. https://doi.org/10.1080/03637751.2016.1201207

Sydnor, E. (2019). *Disrespectful democracy: The psychology of political incivility*. Columbia University Press.

https://doi.org/10.7312/sydn18924
Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(30), 755–769. https://www.jstor.org/stable/3694247
Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.
Thompson, J. B. (2000). *Political scandal: Power and visability in the media age*. Polity Press.
Van 't Riet, J., & Van Stekelenburg, A. (2021). The effects of political incivility on political trust and political participation: A meta-analysis of experimental research. *Human Communication Research*, *48*(2). https://doi.org/10.1093/hcr/hqab022
Vivyan, N., Wagner, M., & Tarlov, J. (2012). Representative misconduct, voter perceptions and accountability: Evidence from the 2009 House of Commons expenses scandal. *Electoral Studies*, *31*(4). https://doi.org/10.1016/J.ELECTSTUD.2012.06.010
Von Sikorski, C., Heiss, R., & Matthes, J. (2020). How political scandals affect the electorate. Tracing the eroding and spillover effects of scandals with a panel study. *Political Psychology*, *41*(3), 549–568. https://doi.org/10.1111/POPS.12638
Walter, A. S. (2014). Negative campaigning in Western Europe: Similar or different? *Political Studies*, *62*(Suppl. 1), 42–60. https://doi.org/10.1111/1467-9248.12084
Walter, A. S. (2021). Introduction: The study of political incivility and its challenges. In A. S. Walter (Ed.), *Political incivility in the parliamentary, electoral and media arena: Crossing boundaries* (pp. 1–16). Routledge.
Walter, A. S., & Vliegenthart, R. (2010). Negative campaigning across different communication channels: Different ball games? *International Journal of Press/Politics*, *15*(4), 441–461. https://doi.org/10.1177/1940161210374122
Warren, M. E. (2017). What kinds of trust does a democracy need? Trust from the perspective of democratic theory. In S. Zmerli & T. W. G. Van der Meer (Eds.), *Handbook on political trust* (pp. 33–52). Edward Elgar Publishing.

About the Authors

Franz Reiter is a PhD candidate at the Department of Communication at the University of Vienna, Austria, where he is a member of the Advertising and Media Psychology Research Group. His main research interests include negative campaigning, political incivility, media content and effects, and political communication. He has published research in international journals such as *Digital Journalism*, *Mass Communication and Society*, and *Social Science Computer Review*.

Jörg Matthes (PhD, University of Zurich) is a professor of communication science at the Department of Communication, University of Vienna, Austria, where he chairs the Advertising and Media Psychology Research Group. His main research interests include advertising, political communication, digital media, and quantitative methods. He has published more than 170 articles in national and international journals and his research activities received numerous awards and honors.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 261–274
https://doi.org/10.17645/pag.v10i4.5677

Article

It's All Relative: Perceptions of (Comparative) Candidate Incivility and Candidate Sympathy in Three Multiparty Elections

Chiara Vargiu

Institute of Political Studies, University of Lausanne, Switzerland; chiara.vargiu@unil.ch

Submitted: 27 April 2022 | Accepted: 4 November 2022 | Published: 30 December 2022

Abstract
While growing attention has been devoted to candidates' use of incivility in campaigns, its role in informing voters' feelings toward candidates is still debated. This study embraces a constructionist perspective on incivility and focuses on the relationship between perceptions of candidate incivility and candidate sympathy. Its contribution is twofold. First, it extends incivility research generalizability by testing the association between voters' perceptions of candidate incivility and candidate sympathy during three election campaigns beyond the US context. Second, it builds upon the notion of incivility as a norm violation and tests the hypothesis that perceptions of a candidate's uncivil behavior are negatively associated with candidate sympathy when this behavior is inappropriate (i.e., it violates injunctive civility norms) and especially when it is uncommon (i.e., it violates descriptive civility norms). These interests are pursued through post-electoral survey data collected in the Netherlands, Germany, and France. Findings show that incivility perceptions can, but not always, correspond to more negative feelings toward candidates. Furthermore, it is the incivility of candidates relative to that of their competitors that really counts for candidate sympathy.

Keywords
candidate incivility; candidate sympathy; France; Germany; incivility perceptions; multiparty systems; survey research; the Netherlands

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

As studies confirm that candidates' sympathy can be decisive for the outcome of elections (Garzia, 2014, 2017), the criteria voters employ to form their feelings toward candidates are of great interest. In this regard, the apparent "coarsening of campaigns" (Stephens et al., 2019) has attracted scholarly attention to candidates' use of incivility—i.e., norm-violating conduct conveying disrespect toward political opponents (Maisel, 2012)—and its effect on candidate sympathy (Druckman et al., 2019; Gervais, 2015; Mutz, 2015). While experimental research generally confirms that incivility lowers candidate sympathy, incivility is still strategically employed (Herbst, 2010), and the electoral success of "roaring candidates" (Maier & Nai, 2020) puts experimental findings into question. Starting from the assumption that voters form their feelings toward candidates based on many considerations, this article investigates the relevance of incivility perceptions in voters' minds by testing their association with candidate sympathy alongside its well-established predictors. I do so through post-election survey data collected after the latest general elections in the Netherlands, Germany, and France. I aim to contribute to incivility research in two ways.

First, I extend incivility research generalizability. Research on political incivility is restricted mainly to the US. However, given the contextual nature of civility

norms, US findings cannot be haphazardly applied to other political systems (Walter, 2021). By interviewing samples of Dutch, German, and French voters, I extend incivility research beyond the US context. Additionally, our knowledge of the effects of incivility on candidate sympathy is primarily based on experimental studies, where participants are usually exposed to an artificial stimulus in the form of a civil or uncivil message from a fictitious politician and then asked to state their feelings toward them (Brooks & Geer, 2007; Druckman et al., 2019; Mutz & Reeves, 2005). While this approach has the advantage of isolating causal mechanisms, it does not account for the fact that voters form their feelings toward candidates based on many considerations. Thus, little is known about the relevance of incivility perceptions in voters' minds. By including perceptions of candidate incivility within traditional models of candidate sympathy, I shed light on their importance in informing candidate sympathy alongside its usual predictors.

Secondly, while scholars agree that incivility is norm-violating conduct, only a few studies have directly drawn from normative theories in their accounts. Based on Muddiman et al.'s (2021) distinction between injunctive civility norms (i.e., what is appropriate when campaigning) and descriptive civility norms (i.e., what is common when campaigning), I argue that candidates' perceived behavior should correspond to more negative feelings toward candidates not only when this behavior deviates from what is perceived as appropriate, but especially when this behavior deviates from what is perceived as common. Thus, this article tests the relationship between candidate sympathy and not only perceived violations of injunctive civility norms (i.e., how uncivil candidates are perceived) but also perceived violations of descriptive civility norms (i.e., how uncivil candidates are perceived relative to their competitors).

Findings show that perceptions of candidate incivility often, but not always, corresponded to more negative feelings toward candidates. There were instances in which perceptions of candidate incivility did not matter for candidate sympathy, and the strength of this relationship varied across candidates and countries. This provides further evidence that incivility is contextual. Additionally, compared to voters' partisan predispositions and their perceptions of candidates' personalities, the role of perceived incivility in informing candidate sympathy was often marginal. Most importantly, results confirmed that the association between perceptions of candidate incivility and candidate sympathy was more consistent and much stronger when candidate incivility perceptions were measured by comparing candidates to one another rather than considering them individually. These results confirm that incivility is a contextual feature of political discourse that could be better understood as a relative rather than an absolute concept.

2. Theoretical Background

2.1. A Constructionist Perspective on Candidate Incivility

While the concept is still debated, political incivility is broadly regarded as a violation of the norms of political discussion (Maisel, 2012). Norms are rules that define what is acceptable behavior among the members of a group (Cialdini & Goldstein, 2004). Therefore, incivility is contextual and whether a behavior is uncivil depends on factors such as the context of the uncivil speech (Roseman et al., 2021), the individual characteristics of its speaker (Muddiman et al., 2021; Sydnor, 2019), and its audience (Kenski et al., 2020). Thus, scholars have advocated for a constructionist approach, emphasizing the role of contextual differences in shaping perceptions and effects of incivility (Jamieson et al., 2017). Following this, I regard incivility as a perceived norm violation that varies according to what is considered normative in each context.

Regarding the nature of the norm violation, scholars distinguish between "public" and "personal" levels of incivility (Muddiman, 2017). The former relates civility to the ideals of deliberative democracies, including behaviors that threaten democratic functioning—e.g., racism, misinformation, or uncompromising conduct (Papacharissi, 2004). Some argue that these behaviors describe intolerant rather than uncivil discourse and deserve attention in their own right (Rossini, 2020). I follow this view and restrict this article's scope to the "personal'' level of incivility. From this perspective, in the same way that during private conversations individuals wish to maintain a "positive face" and expressions of disrespect are considered "face-threatening" (Brown & Levinson, 1987), in public discourse, politicians' disrespectful behavior is considered non-normative (Sobieraj & Berry, 2011). Hence, I equate incivility to disrespectful conduct and focus on behavior such as the use of derogatory language, aggressive speech, or ridicule.

2.2. The Role of Voters' Perceptions of Candidate Incivility Within Models of Candidate Sympathy

Abundant evidence shows that voters develop mental images of politicians based on a set of personal characteristics (Funk, 1996, 1999; Ohr & Oscarsson, 2013). While different trait dimensions have been suggested (Conover, 1981; Funk, 1999; Miller et al., 1986), an influential account has reduced them to four—competence, empathy, leadership, and integrity/honesty (Kinder et al., 1979). Based on this categorization, numerous scholars investigated how perceptions of candidate traits affect feelings toward candidates (e.g., Funk, 1999; Ohr & Oscarsson, 2013; Pancer et al., 1999). In these studies, voters' sympathy toward candidates—usually in the form of feeling thermometers or like-dislike scores—is regressed on voters' partisan predispositions and perceptions of candidate traits. While the relevance of each trait depends on the candidate and the electoral

context (Bittner, 2007; Funk, 1999; Pancer et al., 1999), this empirical work demonstrates that a high score on the feeling thermometer or like–dislike scale is associated with politicians' perceived competence, leadership, empathy, and integrity (Funk, 1999; Ohr & Oscarsson, 2013; Pancer et al., 1999).

Building upon this research, this article tests whether voters' perceptions of candidate incivility are a relevant dimension—alongside partisan predispositions and traits' perceptions—upon which voters base their feelings toward candidates. Considering the widespread concerns over the "coarsening of campaigns" (Stephens et al., 2019), perceptions of a candidate's incivility may have become an important criterion alongside partisanship and perceptions of candidates' personalities. At the same time, it is possible that while citizens generally dislike incivility, incivility perceptions are less relevant for candidate sympathy compared to other considerations. Thus, this article asks whether perceptions of candidate incivility inform candidate sympathy alongside the predictors already identified in the literature. As the role of incivility in models of candidate sympathy remains untested, I keep this question exploratory. Nevertheless, expectations on the direction of the relationship between candidate sympathy and incivility perceptions are formulated in the following subsection.

2.3. The Relationship Between Voters' Perceptions of (Comparative) Candidate Incivility and Candidate Sympathy

Norms are rules that guide behavior, and deviations from such rules are condemned as a threat to social relationships (Cialdini & Goldstein, 2004). Thus, citizens are expected to "punish" behaviors that defy civility norms. US experimental research confirms this view. When exposed to uncivil messages, Americans report lower evaluations of (Gervais, 2015; Maier, 2021; Mutz, 2015) and lower voting intentions for disrespectful candidates (Mutz, 2015). This is true even for partisans, who evaluate uncivil politicians on their side more unfavorably (Druckman et al., 2019; Frimer & Skitka, 2018). A few studies confirm that incivility is punished by the non-US public too. For example, an online experiment on a sample of German voters shows that exposure to incivility from a politician lowers their approval ratings (Maier, 2021). Similarly, Mölders et al. (2017) found that German voters were less willing to vote for disrespectful candidates. Thus, I expect the following:

> H1: There is a direct negative association between voters' perceptions of individual candidate incivility and candidate sympathy.

Despite this theoretical expectation, candidates still employ incivility in their campaigns. To address this paradox, Muddiman et al. (2021) suggest that people may have different opinions on how politicians *should* act and how they *do* act. Drawing from normative theories (Kallgren et al., 2000), they note that scholars have narrowed their definition of incivility to behaviors deviating from what voters approve. However, norms arise not only from what is approved by community members but also from observing what members do. In the former case, scholars refer to *injunctive norms*, i.e., how people *ought to* behave, and in the latter, to *descriptive norms*, i.e., how people do behave (Kallgren et al., 2000). Based on this, Muddiman et al. (2021, p. 13) suggest—but do not directly test—that "if voters feel that uncivil actions are common in campaigns, they may not alter their behaviors even if they do not think the actions are appropriate.''

Building upon this, I suggest that the relationship between candidate incivility perceptions and candidate sympathy could be better understood if we consider not only perceptions of uncivil behavior from single candidates but also from their competitors, as these contribute to determining what is normative. Voters may dislike candidates more not only when they are perceived as violating their injunctive norm of respecting opponents but especially when their injunctively uncivil behavior deviates from what is perceived as common. Hence, this article not only focuses on voters' perceptions of *individual candidate incivility* but also considers voters' perceptions of *comparative candidate incivility*. I propose a comparative measure of candidate incivility perceptions (i.e., how uncivil a candidate is perceived relative to their competitors) and test whether this is more strongly associated with candidate sympathy than an individual measure of candidate incivility perceptions (i.e., how uncivil a single candidate is perceived). I expect the following:

> H2: There is a direct negative association between voters' perceptions of comparative candidate incivility and candidate sympathy.

> H3: Voters' perceptions of comparative (versus individual) candidate incivility are more strongly associated with candidate sympathy.

2.4. The Role of Partisan Sympathy

Considering the well-documented partisan biases in political information acquisition and processing (Campbell et al., 1960), I also consider partisan differences. Voters are motivated reasoners and process information in a way that complies with their partisan predispositions (Taber & Lodge, 2006). At the same time, partisanship, as a relevant social identity, can generate positive feelings for in-parties, and negative feelings for out-parties, thus strengthening intergroup biases (Iyengar & Westwood, 2015). Furthermore, preliminary evidence confirms that incivility is perceived through partisan lenses (Liang & Zhang, 2021). Thus, I expect perceptions of incivility to matter more in negatively evaluating candidates far away from oneself than those closer. I hypothesize the following:

H4: The negative association between voters' perceptions of (individual and comparative) candidate incivility and candidate sympathy is weaker for respondents with more positive feelings toward the candidate's party, while it is stronger for respondents with more negative feelings toward the candidate's party.

2.5. The Role of Candidate Differences: Gender, Populism, Incumbency

Following the constructionist approach, scholars have also suggested that the same behavior may be perceived differently depending on the characteristics of the person engaging in it (Jamieson et al., 2017; Muddiman et al., 2021; Sydnor, 2019). Three features seem particularly relevant. First, the stereotype content model predicts that women are associated with communal traits (e.g., warmth), while men are associated with agentic traits (e.g., dominance; Fiske et al., 2002). As incivility goes against the expectations of women being kind, perceptions of incivility by female candidates should lead to stronger backlashes. Secondly, populist candidates are known for their transgressive style of campaigning (Moffitt & Tormey, 2013); thus, people should respond less negatively to perceptions of incivility from populists as these align with what is expected of them (Nai et al., 2022). Finally, incumbents prefer to rely on their political records to promote themselves rather than attack and risk a backlash (Nai, 2018). Because being uncivil is rarer for incumbents, they pay a higher price when they do so (Fridkin & Kenney, 2011). Based on this, I hypothesize the following:

H5: The negative association between voters' perceptions of (individual and comparative) candidate incivility and candidate sympathy is stronger for (a) female (versus male) candidates, (b) populist (versus non-populist) candidates, and (c) incumbents (versus challengers).

3. Data and Methods

3.1. Case Selection

The hypotheses are tested on Dutch, German, and French samples. Doing so adds to the study of incivility in multiparty systems. Here, the necessity to engage in cross-partisan compromise may not only discourage the use of incivility against potential coalition partners but also strengthen individuals' negative reactions to it. These features set multiparty democracies apart in using and processing incivility, making them relevant cases to extend our knowledge on incivility. Additionally, these cases are characterized by crucial variations in terms of electoral competition, including an entirely proportional system (the Netherlands), a mixed system (Germany), and a semi-presidential majoritarian system (France). As in contexts with high party fragmentation incivility more directly threatens negotiations for government formation between parties, this selection allows me to assess the robustness of my findings across party structures.

3.2. Sample

Online samples of the Dutch (from 22–03–2021 to 05–04–2021; *N* = 1,007), German (from 28–09–2021 to 12–10–2021; *N* = 999), and French (from 06–05–2022 to 19–05–2022; *N* = 1,246) populations eligible to vote were collected within three larger post-election surveys distributed through private survey companies (Kantar in the Netherlands, Dynata in Germany and France). Inattentive respondents (i.e., respondents whose completion time was less than half of the median of the country's sample) and straightliners (i.e., respondents whose responses on the candidate incivility perceptions and the candidates like–dislike batteries had a standard deviation of zero) were excluded. This resulted in three final samples of *n* = 898 (NL), *n* = 804 (DE), and *n* = 1,063 (FR). As a robustness check, main analyses are replicated with the inclusion of straightliners, providing consistent results (see Appendix C of the Supplementary File, Table C1). The online samples do not represent the general voting population but employ stratification quotas for gender, age, and macro-region of residence (see Supplementary File, Table A1).

Respondents were interviewed about a selection of six candidates in the Netherlands and Germany and eight in France (see Supplementary File, Table A2). While it was not feasible to interview respondents about the entire population of candidates, this selection covered almost 70% of the electoral preferences in the Netherlands (Kiesraad, 2021), more than 85% in Germany (The Federal Returning Officer, 2021), and almost 90% in France (AFP, 2022), and included representatives of all relevant party families in the West European context.

3.3. Candidate Sympathy

Candidate sympathy was measured through like-dislike scores, widely employed in electoral research to gauge voters' overall feelings toward political leaders (Garzia, 2017) and in incivility research to test the effects of incivility on sponsor perceptions (e.g., Druckman et al., 2019; Gervais, 2015; Mutz, 2015). Concretely, respondents were asked to assign a score from 0 (*dislike*) to 10 (*like*) to each candidate, and the *candidate sympathy* scale was used as the dependent variable in all models (for summary statistics, see Supplementary File, Table A3).

3.4. Candidate Incivility Perceptions

3.4.1. Individual Candidate Incivility Perceptions

Incivility can take many forms, and this article focuses on behaviors commonly employed in previous research (e.g., Muddiman, 2021; Mutz, 2015; Otto et al., 2020).

Respondents rated candidates on a scale from 0 (*not at all*) to 10 (*very much*) on the extent to which during the campaign they engaged in three kinds of uncivil behaviors: (a) They used insulting or derogatory language; (b) they employed forms of emotionalized speech, such as by shouting; and (c) they ridiculed their opponents (see Supplementary File, Table A4 for the full question text, and Table A5 for summary statistics). For each candidate, a scale of *individual candidate incivility perceptions* was computed by calculating the means of the three incivility items. Summary statistics, including reliability scores (all above 0.80), are summarized in the Supplementary File (Table A6).

3.4.2. Comparative Candidate Incivility Perceptions

A measure of *comparative candidate incivility perceptions* was computed by subtracting from candidates' individual incivility perceptions the average incivility perceptions of their competitors. For example, Rutte's comparative incivility perceptions score was built by subtracting the average individual incivility perceptions scores voters assigned to the other five candidates from the individual incivility perceptions' score they assigned to Rutte. This resulted in a scale ranging from −10 to +10, recoded to range from 0 to 10 to ease comparisons. A score of five means that Rutte is perceived *as uncivil as his competitors*; increasing scores above five mean that Rutte is perceived as *increasingly more uncivil than his competitors*; decreasing scores below five mean that Rutte is perceived as *increasingly less uncivil than his competitors* (for summary statistics, see Supplementary File, Table A7).

3.5. Other Predictors of Candidate Sympathy

Models include the usual predictors of candidate sympathy as identified in the literature summarized above, i.e., partisan predispositions and perceptions of candidate traits. Respondents indicated on a scale from 0 (*dislike*) to 10 (*like*) how much they liked each candidate's party (for summary statistics, see Supplementary File, Table A8). This measure of *partisan sympathy* was included in the models as a control variable and then to test its interaction with perceptions of candidate incivility. Regarding perceptions of candidate traits, I focus on the four main criteria identified by Kinder et al. (1979): empathy, honesty, competence, and leadership skills. Respondents were asked to rate candidates on these attributes on a scale from 1 to 5. The *leadership*, *competence*, *empathy*, and *honesty perceptions scales* were computed (see Supplementary File, Table A9) and included in the models.

3.6. Candidates' Characteristics

Candidates were categorized based on gender, populism, and incumbency status. Populist candidates were selected based on the categorization by Rooduijn et al. (2019), while incumbency was defined as candidates who held a position within the government before the elections. Three nominal levels variables (Gender: 0 = *male*, 1 = *female*; Populism: 0 = *non-populist*, 1 = *populist*; Incumbency: 0 = *challenger*, 1 = *incumbent*) were computed (see Supplementary File, Table A10).

3.7. Analysis Plan and Modelling Strategy

After presenting the results of a descriptive analysis exploring how much incivility was perceived during the three elections, I test the bivariate relationship between individual and comparative candidate incivility perceptions and candidate sympathy. Then, I formally test H1, H2, and H3 following previous work on candidate authenticity by Stiers et al. (2021). In their account of trait authenticity, they test the relevance of candidates' perceived authenticity by regressing candidate sympathy on its traditional predictors—partisan predispositions and candidate traits—with the addition of their newly developed scale of trait authenticity. Following this modelling strategy, I conduct two OLS multiple regression models per country predicting candidate sympathy from perceptions of individual (M1) and comparative (M2) candidate incivility, including socio-demographics (age, gender, and education) and perceptions of candidate traits and partisan sympathy. I subsequently test H4 by including an interaction term (candidate incivility perceptions * party sympathy) to M1 and M2. Finally, I test H5 by pooling data from each country and running six additional OLS regression models predicting candidate sympathy from individual and comparative candidate incivility perceptions and their respective interaction with candidates' gender, populism, and incumbency status.

All models are run on a stacked dataset with voter–candidate as the unit of analysis, including robust and clustered standard errors. Voters' characteristics are included following the procedure by van der Eijk et al. (2006). First, each individual measure of candidate sympathy was regressed on gender, age, and education in the unstacked data matrix. Then, the predicted values (y-hats) of each separate regression were included in the stacked data matrix. As a robustness check, models are replicated for each candidate separately, providing consistent results (see Supplementary File, Tables B2–B7).

4. Results

4.1. Descriptive Analysis

Figure 1 shows the means of individual and comparative incivility perceptions. In Germany and the Netherlands, perceptions of individual incivility are relatively low but with large differences across candidates. Notably, right-wing populist candidates are perceived as the most uncivil. In the Netherlands, Wilders (M = 6.13, SD = 2.83) is perceived more than twice as uncivil as the second

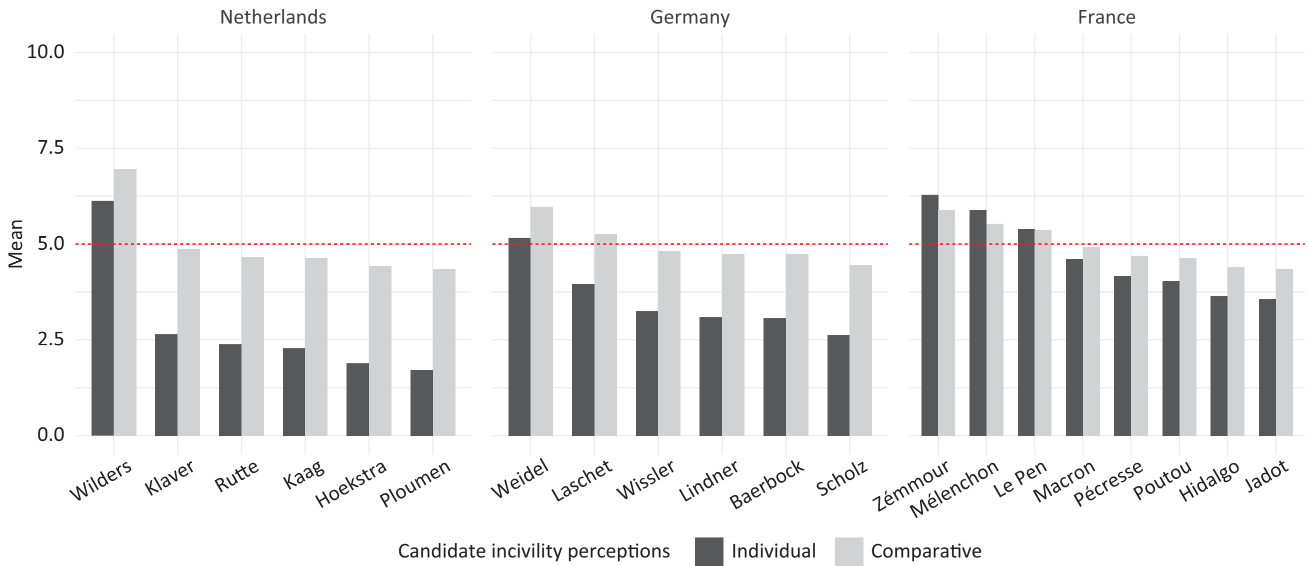

Figure 1. Means of individual and comparative candidate incivility perceptions.

most uncivil candidate (Klaver: *M* = 2.65, *SD* = 2.45) and more than four times as uncivil as the least uncivil candidate (Ploumen: *M* = 1.71, *SD* = 2.08). In Germany, perceptions of individual incivility are more evenly distributed, ranging from 2.63 (*SD* = 2.56) for the least uncivil candidate (Scholz) to 5.16 (*SD* = 3.18) for the most uncivil candidate (Weidel), with a gap of around one unit from the second most uncivil candidate (Laschet: *M* = 3.97, *SD* = 2.76). French candidates are perceived as relatively more uncivil; individual incivility perceptions range between 3.56 (Jadot: *SD* = 2.68) and 6.29 (Zemmour: *SD* = 3.04). While populist leaders are perceived as the most uncivil in France too, this difference is less pronounced, with less than a one-point distance between the least uncivil populist candidate (Le Pen: *M* = 5.39, *SD* = 2.92) and the first most uncivil non-populist candidate (Macron: *M* = 4.61, *SD* = 2.99).

Regarding perceptions of comparative incivility, scores are close to the middle of the scale, suggesting that, on average, voters tend to perceive comparable levels of incivility across candidates. Again, populists stand out. With the only exception of Germany—where conservative candidate Laschet scored on average as relatively more uncivil than his competitors, and left-wing populist candidate Wissler is on average perceived less uncivil than her competitors—populist candidates are the only ones with comparative incivility perceptions scores above the middle of the scale in all countries.

4.2. Bivariate Analysis

Table 1 presents the results of a bivariate analysis testing the relationship between, on the one hand, individual and comparative candidate incivility perceptions and, on the other, candidate sympathy. Higher scores on the individual incivility perceptions scale always correspond to lower scores on the candidate sympathy scale. Dutch candidates show moderate to strong correlations, while associations are relatively weaker for French and German candidates. Turning to comparative incivility

Table 1. Pearson's *r* correlation coefficients and significance levels for the relationship between individual and comparative candidate incivility perceptions and candidate sympathy.

The Netherlands			Germany			France		
Incivility Perceptions			Incivility Perceptions			Incivility Perceptions		
	Ind.	Comp.		Ind.	Comp.		Ind.	Comp.
Wilders	−0.56***	−0.63 ***	Weidel	−0.38***	−.56 ***	Zemmour	−0.36***	−0.54***
Rutte	−0.52***	−0.55***	Laschet	−0.18***	−.39 ***	Le Pen	−0.36***	−0.40***
Hoekstra	−0.24***	−0.24***	Lindner	−0.15***	−.29 ***	Macron	−0.39***	−0.45***
Kaag	−0.49***	−0.55***	Scholz	−0.16***	−.25 ***	Pécresse	−0.19***	−0.39***
Ploumen	−0.34***	−0.27***	Baerbock	−0.29***	−.46 ***	Hidalgo	−0.17***	−0.34***
Klaver	−0.39***	−0.47***	Wissler	−0.12**	−.33 ***	Jadot	−0.15***	−0.32***
						Mélenchon	−0.36***	−0.48***
						Poutou	−0.26***	−0.39***

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

perceptions, candidate sympathy is almost always more strongly negatively associated with comparative rather than individual incivility perceptions. This is especially evident in Germany and France, where the relationship is, in many cases, twice as strong. This is in line with H1, H2, and H3; perceptions of candidate incivility are negatively associated with candidate sympathy, and this association is especially strong when incivility perceptions are measured comparatively.

4.3. Multiple Regression Analysis

While results from the bivariate analysis are consistent with H1, H2, and H3, the relationship between candidate incivility perceptions and candidate sympathy needs to be assessed within more demanding models. Figure 2 presents point estimates and 95% confidence intervals for individual (M1) and comparative (M2) candidate incivility perceptions from two OLS multiple regression models, including socio-demographics (gender, age, education), partisan sympathy, and candidate trait perceptions. Figure 3 depicts the marginal effects of individual (M1) and comparative (M2) incivility perceptions on candidate sympathy (full models are summarized in the Supplementary File, Table B1). Table 2 presents the proportion of explained variance in M1 and M2 in each country, in comparison with the base model excluding the incivility perceptions measures (M0). M1 and M2 are replicated for the three forms of incivility (insults, negative emotions, and sarcasm), providing similar results (see Supplementary File, Tables C2, C3, and C4).

Starting with M1, there is a negative statistically significant association between individual incivility perceptions and candidate sympathy in the Netherlands and France but not Germany. In the Netherlands, from the lowest to the highest individual incivility perceptions score, there is a significant decrease in candidate sympathy of 11%. In France, this decrease amounts to only 4%. Comparing regression coefficients across predictors, their size is much smaller for individual incivility perceptions than it is for other predictors. Hence, even if there are significant negative associations

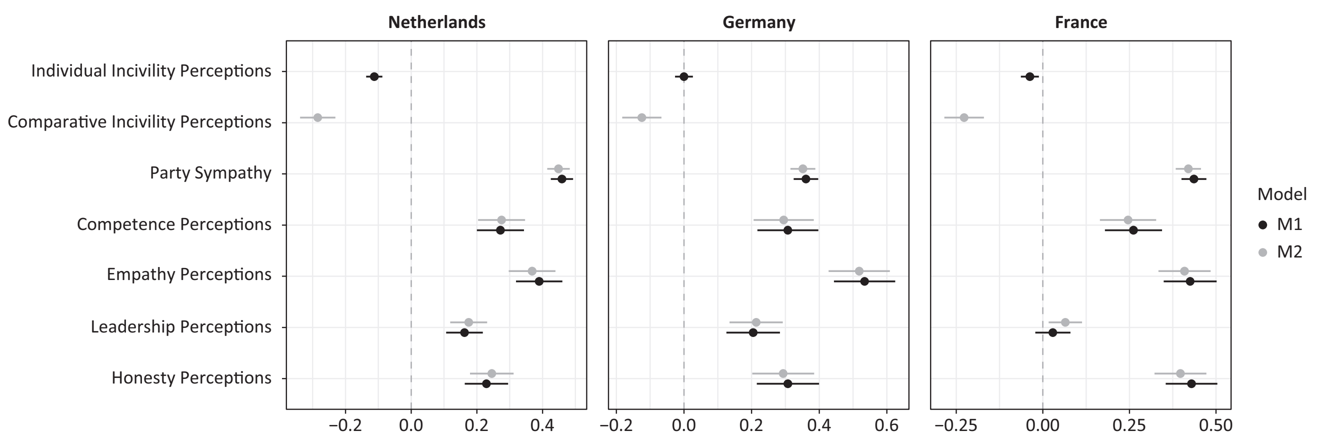

Figure 2. Results of OLS multiple regression models predicting candidate sympathy from individual (M1) and comparative (M2) incivility perceptions.

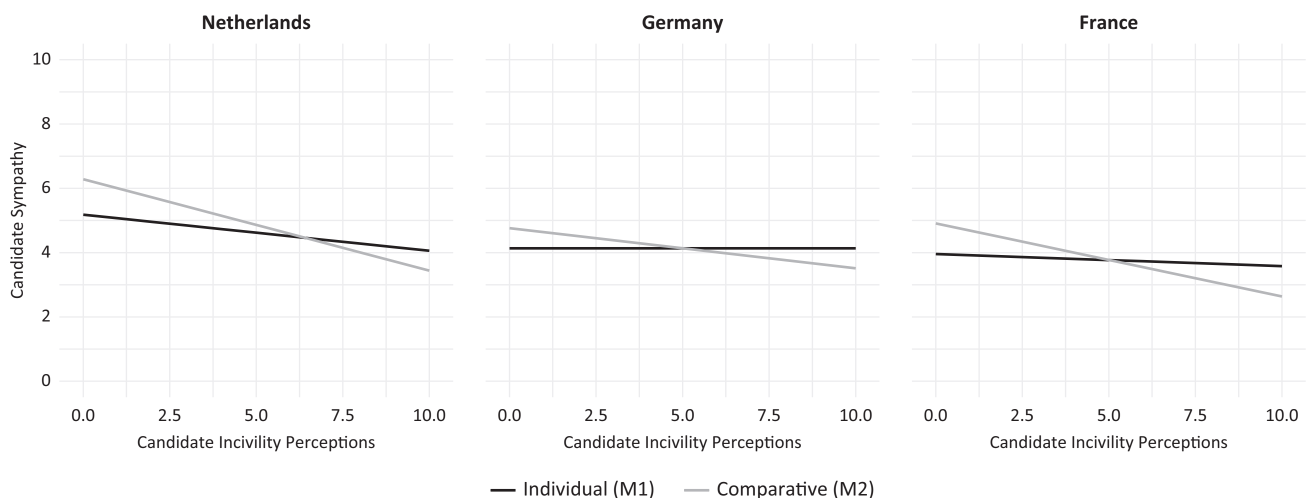

Figure 3. Marginal effects of individual (M1) and comparative (M2) candidate incivility perceptions on candidate sympathy. Note: Results from OLS regression M1 and M2.

Table 2. Adjusted R^2 from OLS regression models predicting candidate sympathy.

	Adjusted R^2		
	M0	M1	M2
The Netherlands	0.73	0.74	0.74
Germany	0.71	0.71	0.71
France	0.65	0.65	0.66

between individual incivility perceptions and candidate sympathy—H1 is confirmed in the Netherlands and France—the role of individual candidate incivility perceptions is marginal compared to those of the other considerations.

Turning to M2, the association between incivility perceptions and candidate incivility is more consistent across countries and much stronger when perceptions of incivility are measured comparatively. Comparative perceptions of incivility are significantly associated with lower levels of candidate sympathy in all three samples, and their regression coefficients are similar in range to other predictors. An additional unit in the comparative incivility perceptions scale corresponds to a significant decrease in candidate sympathy of 28% (vs. 11% in M1) in the Netherlands, 12% in Germany (vs. a non-significant association in M1), and 23% (vs. 4% in M1) in France. These results confirm H2 and H3; higher levels of comparative incivility perceptions correspond to lower levels of candidate sympathy, and the negative association between incivility perceptions and candidate sympathy is much stronger when incivility perceptions are measured comparatively.

Finally, the inclusion of candidate incivility perceptions does not improve the models' explained variances compared to the base models (see Table 2).

4.4. The Role of Partisan Sympathy

I now test for partisan differences by including an interaction term between party sympathy and incivility perceptions to M1 and M2. This resulted in two OLS regression models per country (M1.INT and M2.INT). Figure 4 presents point estimates and confidence intervals of focal independent variables (for full numerical results, see Supplementary File, Table B1). In the Netherlands, there are no significant differences in the association between individual and comparative incivility perceptions and candidate sympathy at levels of party sympathy. On the contrary, the interaction term is significant for both measures of incivility perceptions in Germany (M1.INT: $b = -0.01$, $p < 0.05$; M2.INT: $b = -0.03$, $p < 0.001$) and only for individual incivility perceptions in France (M1.INT: $b = -0.01$, $p < 0.05$).

Figure 5 graphically presents the slopes of individual and comparative candidate incivility perceptions along the party sympathy scale and the range of values of party sympathy in which these slopes are significant versus insignificant. Table 3 summarizes the slopes of individual and comparative incivility perceptions at levels of party sympathy, calculated at one standard deviation below and above the mean. Starting with M1.INT, at higher levels of partisan sympathy, an additional unit in the individual candidate incivility perceptions scale corresponded to a decrease in the candidate sympathy scale of 3% in Germany and 7% in France. This compares to an insignificant association at lower levels of partisan sympathy and, interestingly, to an increase of 2% in the candidate sympathy scale for every additional unit of individual incivility perceptions in Germany.

Turning to M2.INT, the interaction between comparative incivility perceptions and party sympathy was significant only in Germany, where at higher levels of partisan sympathy, an additional unit in the comparative incivility perceptions scale corresponded to a 23% decrease in candidate sympathy, compared to only a 6% decrease at lower levels of partisan sympathy. While there were no significant differences at different levels

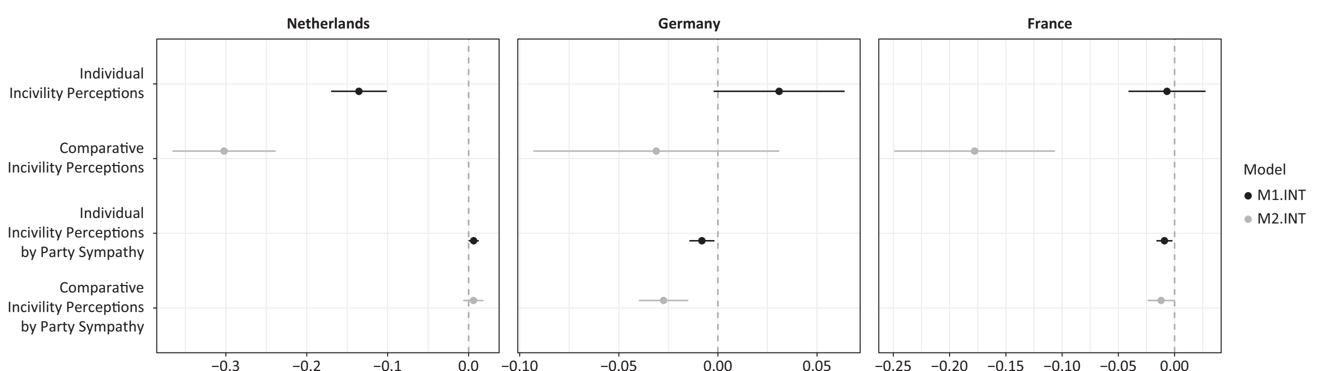

Figure 4. Results of OLS multiple regression models predicting candidate sympathy from individual (M1.INT) and comparative (M2.INT) candidate incivility perceptions and their interaction with party sympathy.

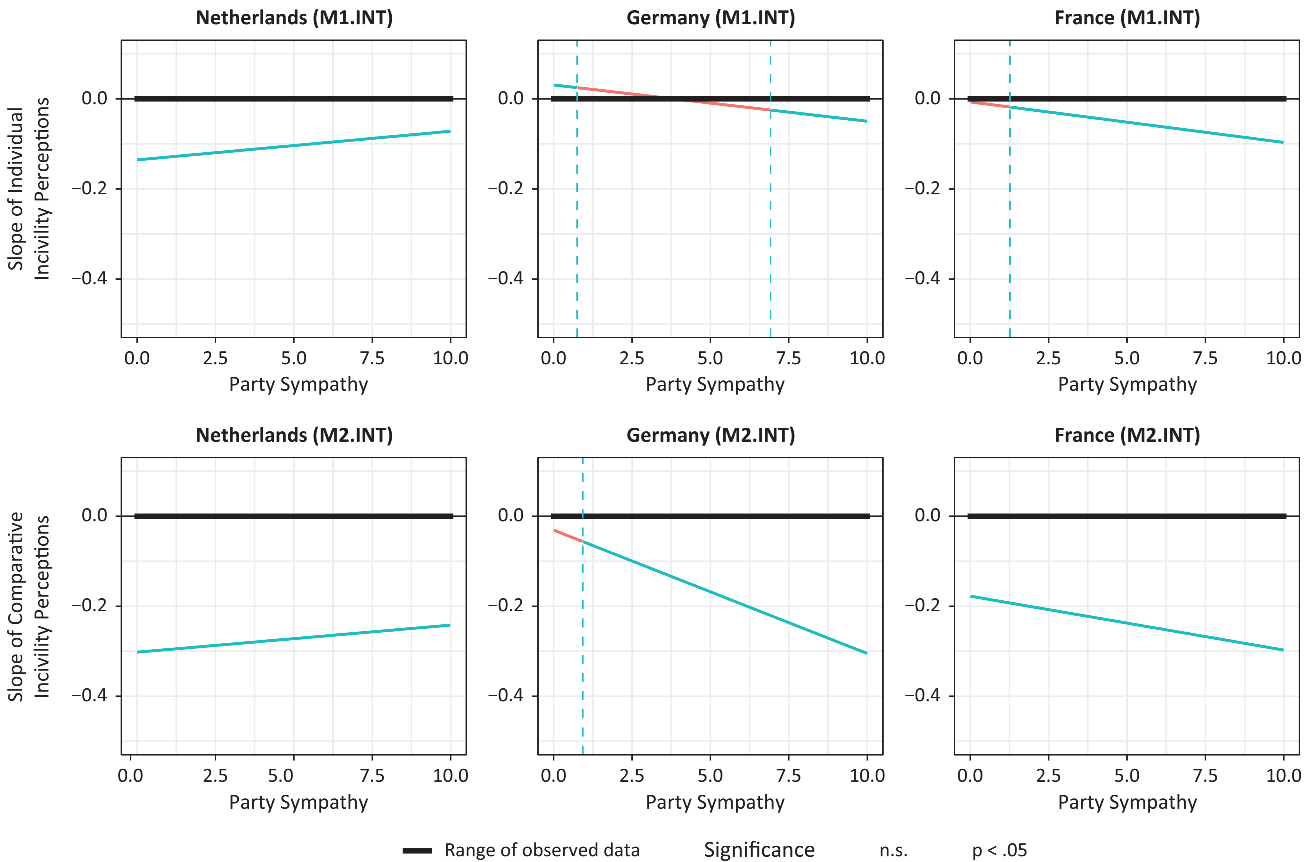

Figure 5. Johnson-Neyman plots. Notes: Results from OLS regression M1.INT and M2.INT; $p < 0.05$.

of party sympathy in France, a similar pattern can be identified. As shown in Figure 5, the slope of comparative incivility perceptions increases as levels of partisan sympathy also increase. Overall, these results disprove H4: When the interaction between candidate incivility perceptions and candidate sympathy was significant, perceptions of candidate incivility mattered more for individuals with more positive rather than negative partisan sympathy.

4.5. The Role of Candidate Characteristics

Finally, I have run six additional OLS regression models predicting candidate sympathy from individual and comparative candidate incivility perceptions and their respective interaction with candidates' gender, populism, and incumbency status. Table 4 summarizes the coefficients of these interaction terms in each model. Except for the interaction between populism and individual incivility perceptions, all interaction coefficients are significant.

Figure 6 shows the marginal effects of individual and comparative incivility perceptions at the levels of each moderator. The relationship between incivility perceptions and candidate sympathy remains significantly negative for all candidates, regardless of their gender, whether they are populist, and their incumbency status. Differences only arise in the magnitude of this relationship which is significantly more strongly negative for male candidates (H5a is not confirmed), non-populist candidates (only in relation to comparative

Table 3. Slopes of individual (M1) and comparative (M2) candidate incivility perceptions at values of party sympathy.

	Netherlands		Germany		France	
Party Sympathy	M1	M2	M1	M2	M1	M2
Mean − 1 SD	−0.12*** (0.01)	−0.29*** (0.02)	0.02* (0.01)	−0.06* (0.03)	−0.01 (0.01)	−0.19*** (0.01)
Mean	−0.11*** (0.01)	−0.27*** (0.03)	0.00 (0.01)	−0.15*** (0.03)	−0.04*** (0.01)	−0.22*** (0.01)
Mean + 1 SD	−0.09*** (0.01)	−0.26*** (0.03)	−0.03* (0.01)	−0.23*** (0.03)	−0.07*** (0.01)	−0.26*** (0.01)

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 4. Slopes of individual (M1) and comparative (M2) candidate incivility perceptions at levels of gender, populism, and incumbency.

	Gender			Populism			Incumbency		
	b	*se*	*b*	*b*	*b*	*p*	*b*	*se*	*p*
M1	−0.02	0.01	*	−0.02	0.01		−0.04	0.01	**
M2	−0.07	0.02	**	−0.09	0.02	***	−0.08	0.03	**

Notes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

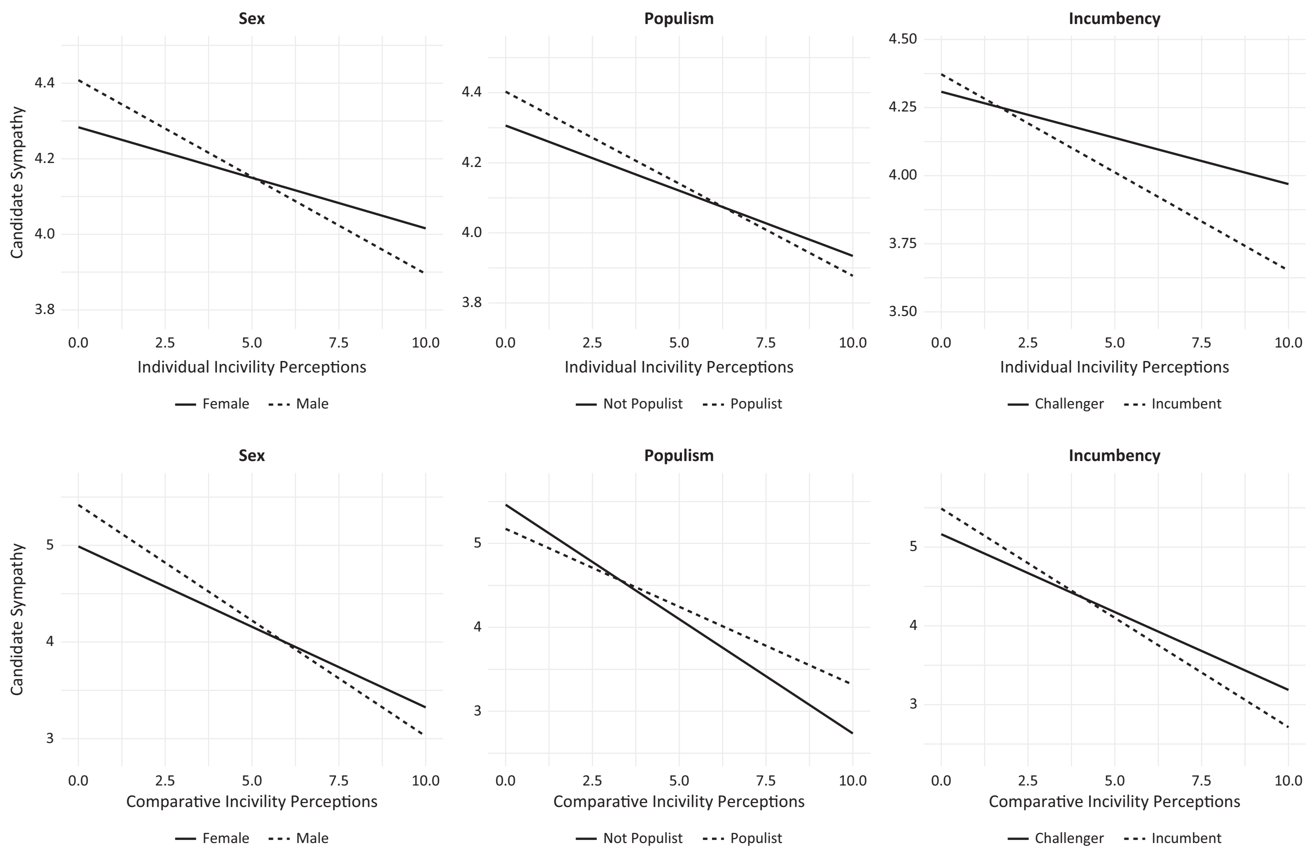

Figure 6. Marginal effects. Note: Results from OLS regression M1 and M2.

incivility; H5b is partially confirmed), and incumbents (H5c is confirmed).

5. Limitations

Before discussing the results of this study, some of its limitations must be acknowledged. Firstly, this study is correlational. While experimental evidence has demonstrated a causal relationship from incivility to candidate sympathy, the opposite may also be true. As partisan feelings shape perceptions of incivility (Liang & Zhang, 2021), I cannot exclude the possibility that the hypothesized relationships also go in the opposite direction. Additionally, it cannot be excluded that perceptions of incivility impact voters' impressions of candidate traits and, through them, affect candidate sympathy. Preliminary evidence suggests that incivility can lower voting intentions by lowering communion judgments (e.g., politicians' perceived friendliness), while it does not affect agency judgment (e.g., politicians' perceived confidence; Mölders et al., 2017). Hence, future research should causally test the interplay between incivility perceptions, candidate traits, and candidate sympathy.

Secondly, data were collected in the aftermath of the elections. Research demonstrates that losing an election can negatively affect voters' judgments. For instance, losers show higher dissatisfaction with democracy (Hansen et al., 2019) and more negative opinions about elections' integrity (Cantú & García-Ponce, 2015). Thus, losers may perceive greater incivility than winners. It is also debatable whether voters can recall how uncivil candidates had been once elections are over. Therefore, this study should be replicated with pre-electoral data. This would provide a measure of perceptions of candidate incivility that is not colored by knowing who lost or won or by memory impairments.

Finally, results showed variations based on candidates' characteristics. However, these differences must be carefully considered on account of the small number of candidates and the fact that I could not control for the confounding effect of multiple candidate characteristics (e.g., female populist candidates vs. female non-populist candidates). Further research should replicate these analyses with a larger sample of candidates, and experimental studies should attempt to isolate the effects of candidate characteristics.

6. Discussion

This study embraced a constructionist definition of incivility and focused on perceptions of candidate incivility and their relationship to candidate sympathy during the last general elections in the Netherlands, Germany, and France. Firstly, results showed that perceptions of incivility from a candidate were associated with more negative feelings toward that candidate. While these results corroborated US experimental findings, the magnitude of this negative relationship was relatively weak compared to usual predictors of candidate sympathy. This is important to note as incivility not only has negative consequences but also can be entertaining and attention-grabbing (Borah, 2014; Mutz & Reeves, 2005), thus newsworthy (Muddiman, 2013). This could partially explain the paradox whereby candidates go uncivil even though most people dislike it. Candidates may resort to incivility despite its potential to lower likeability, as this may be counterbalanced by the positive consequences of an increase in visibility.

Secondly, results confirmed that voters have more negative feelings toward candidates if their perceived incivility deviates not only from what is appropriate but especially from what is common. As predicted, the association between candidate sympathy and incivility perceptions was stronger and more consistent when perceptions of candidate incivility were measured in a relative fashion. These results highlight the importance of distinguishing between injunctive and descriptive norm violations, which should take center stage in future studies. They also suggest that incivility could be better understood as a relative rather than an absolute concept. This is in line with the idea of incivility as a contextual feature of political discourse, which not only depends on one's behavior but also on the behavior of relevant others. Extending this logic to intergroup dynamics, future research could investigate people's reactions to intergroup perceptions of candidate incivility, i.e., perceptions of opposition candidates net of perceptions of favorite candidates. Since partisans generally perceive a higher level of incivility from their outgroups than from their ingroups (Liang & Zhang, 2021; Muddiman, 2021), the relative dynamics of incivility may be stronger when comparative perceptions are measured in a partisan fashion.

Thirdly, while results broadly hold across countries, there are some variations. Dutch respondents showed the strongest negative association between perceptions of candidate incivility and candidate sympathy, which may highlight significant contextual differences. In the Netherlands, the large number of parties makes the political landscape much more fragmented, and the need to form coalitions has pushed political elites into "a style of political accommodation rather than political competition" (Bovens & Wille, 2008, p. 296). In this context, voters may regard incivility as a greater threat to consensual politics, thus punishing it more. Additionally, the three countries differ in the amount of perceived incivility. In Germany, where voters showed the lowest levels of perceived candidate incivility, incivility may be less salient in voters' minds. In France, where voters showed the highest levels of perceived candidate incivility, incivility may be seen as the norm, thus, less relevant. These findings underlie the need to consider differences in party systems and political cultures.

Finally, results showed that candidates' and voters' characteristics matter too. First, there were differences in levels of partisan sympathy. Contrary to expectations, individuals with more positive (rather than negative) partisan feelings showed a stronger negative association between perceptions of candidate incivility and candidate sympathy. This unexpected finding could be related to floor effects. At high levels of partisan dislike, respondents also showed very low levels of candidate sympathy. Hence, they could only move so far in the candidate sympathy scale at levels of incivility perceptions. Nevertheless, this result needs further investigation as it is consistent with the so-called "black-sheep effect" (Reese et al., 2013). Stronger partisans may be more critical toward ingroup deviants to preserve a positive group identity. Secondly, while the association between candidate incivility perceptions and candidate sympathy was consistently negative across candidate types, it was stronger for male candidates, non-populist, and incumbents. These results must be assessed considering the limitations mentioned above. Yet, they provide further evidence that incivility is contextual.

Acknowledgments

I would like to thank the anonymous reviewers for all their valuable comments and suggestions, which greatly helped me improve the quality of the manuscript.

Conflict of Interests

The author declares no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

AFP. (2022). *Présidentielle 2022: Les résultats en direct* [Presidential elections 2022. Live results]. https://interactive.afp.com/graphics/Presidentielle-2022-les-resultats-en-direct_622

Bittner, A. (2007, May 30–June 1). *Competence or charisma? The effects of measurement on why some leaders' traits matter more than others* [Paper presentation]. Canadian Political Science Annual Meeting, Saskatoon, Canada.

Borah, P. (2014). Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication Research*, *41*(6), 809–827. https://doi.org/10.1177/0093650212449353

Bovens, M., & Wille, A. (2008). Deciphering the Dutch drop: Ten explanations for decreasing political trust in the Netherlands. *International Review of Administrative Sciences*, *74*(2), 283–305. https://doi.org/10.1177/0020852308091135

Brooks, J. D., & Geer, J. G. (2007). Beyond negativity: The effects of incivility on the electorate. *American Journal of Political Science*, *51*(1), 1–16. https://www.jstor.org/stable/4122902

Brown, P., & Levinson, C. S. (1987). *Politeness: Some universals language usage*. Cambridge University Press.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. Wiley.

Cantú, F., & García-Ponce, O. (2015). Partisan losers' effects: Perceptions of electoral integrity in Mexico. *Electoral Studies*, *39*, 1–14. https://doi.org/10.1016/J.ELECTSTUD.2015.03.004

Cialdini, R., & Goldstein, N. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621. https://doi.org/10.1146/ANNUREV.PSYCH.55.090902.142015

Conover, P. J. (1981). Political cues and the perception of candidates. *American Political Research*, *9*(4), 427–448. https://doi.org/10.1177/1532673X8100900403

Druckman, J. N., Gubitz, S. R., Levendusky, M. S., & Lloyd, A. M. (2019). How incivility on partisan media (de)polarizes the electorate. *Journal of Politics*, *81*(1), 291–295. https://doi.org/10.1086/699912

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022–3514.82.6.878

Fridkin, K. L., & Kenney, P. (2011). Variability in citizens' reactions to different types of negative campaigns. *American Journal of Political Science*, *55*(2), 307–325. https://doi.org/10.1111/j.1540–5907.2010.00494.x

Frimer, J. A., & Skitka, L. J. (2018). The Montagu principle: Incivility decreases politicians' public approval, even with their political base. *Journal of Personality and Social Psychology*, *115*(5), 845–866. https://doi.org/10.1037/pspi0000140

Funk, C. L. (1996). Understanding trait inferences in candidate images. In M. X. Delli (Ed.), *Research in micropolitics* (Vol. 5, pp. 97–123). JAI Press.

Funk, C. L. (1999). Bringing the candidate into models of candidate evaluation. *Journal of Politics*, *61*(3), 700–720. https://doi.org/10.2307/2647824

Garzia, D. (2014). *Personalization of politics and electoral change*. Palgrave Macmillan.

Garzia, D. (2017). Voter evaluations of candidates and party leaders. In K. Arzheimer, J. Evans, & M. Lewis-Beck (Eds.), *The SAGE handbook of electoral behaviour* (pp. 634–653). SAGE.

Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology and Politics*, *12*(2), 167–185. https://doi.org/10.1080/19331681.2014.997416

Hansen, S. W., Klemmensen, R., & Serritzlew, S. (2019). Losers lose more than winners win: Asymmetrical effects of winning and losing in elections. *European Journal of Political Research*, *58*(4), 1172–1190. https://doi.org/10.1111/1475–6765.12329

Herbst, S. (2010). *Rude democracy: Civility and incivility in American politics*. Temple University Press.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690–707. https://doi.org/10.1111/ajps.12152

Jamieson, K., Volinsky, A., Weitz, I., & Kenski, K. (2017). Political uses and abuses of civility and incivility. In K. Kenski & K. H. Jamieson (Eds.), *The Oxford handbook of political communication* (pp. 205–218). Oxford University Press.

Kallgren, C., Reno, R., & Cialdini, R. (2000). A focus theory of normative conduct: When norms do and do not affect behavior. *Personality and Social Psychology Bulletin*, *26*(8), 1002–1012. https://doi.org/10.1177/01461672002610009

Kenski, K., Coe, K., & Rains, S. A. (2020). Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research*, *47*(6), 795–814. https://doi.org/10.1177/0093650217699933

Kiesraad. (2021). *Officiële uitslag Tweede Kamerverkiezing 17 maart 2021* [Official results of the parliamentary elections of March 17, 2021]. https://www.kiesraad.nl/actueel/nieuws/2021/03/26/officiele-uitslag-tweede-kamerverkiezing-17-maart-2021

Kinder, D. R., Abelson, R. P., & Fiske, S. T. (1979). *Developmental research on candidate instrumentation: Results and recommendation*. American National Election Studies. https://electionstudies.org/wp-content/uploads/2018/07/nes002237.pdf

Liang, H., & Zhang, X. (2021). Partisan bias of perceived incivility and its political consequences: Evidence from survey experiments in Hong Kong. *Journal*

of Communication, *71*(3), 357–379. https://doi.org/10.1093/JOC/JQAB008

Maier, J. (2021). Will fouls in politics be punished? An experiment on the effect of incivility on the support for the sender and for the political system. In A. S. Walter (Ed.), *Political incivility in the parliamentary, electoral and media arena* (pp. 200–218). Routledge.

Maier, J., & Nai, A. (2020). Roaring candidates in the spotlight: Campaign negativity, emotions, and media coverage in 107 national elections. *The International Journal of Press/Politics*, *25*(4), 576–606. https://doi.org/10.1177/1940161220919093

Maisel, L. S. (2012). The negative consequences of uncivil political discourse. *Political Science and Politics*, *45*(3), 405–411. https://doi.org/10.1017/S1049096512000467

Miller, A. H., Wattenberg, M. P., & Malanchuk, O. (1986). Schematic assessments of presidential candidates. *American Political Science Review*, *80*(2), 521–540. https://doi.org/10.2307/1958272

Moffitt, B., & Tormey, S. (2013). Rethinking populism: Politics, mediatisation and political style. *Political Studies*, *62*(2), 381–397. https://doi.org/10.1111/1467–9248.12032

Mölders, C., van Quaquebeke, N., & Paladino, M. P. (2017). Consequences of politicians' disrespectful communication depend on social judgment dimensions and voters' moral identity. *Political Psychology*, *38*(1), 119–135. https://doi.org/10.1111/pops.12311

Muddiman, A. (2013). *The instability of incivility: How news frames and citizen perceptions shape conflict in American politics* [Doctoral dissertation, University of Texas]. Texas ScholarWorks. https://repositories.lib.utexas.edu/handle/2152/23311

Muddiman, A. (2017). Personal and public levels of political incivility. *International Journal of Communication*, *11*, 3182–3202. https://doi.org/1932–8036/20170005

Muddiman, A. (2021). Conservatives and incivility. In S. E. Jarvis (Ed.), *Conservative political communication* (pp. 119–136). Routledge.

Muddiman, A., Flores, L., & Boyce, B. (2021). Descriptive and injunctive incivility norms in political campaigns: Differences across behavior type, candidate gender, and candidate party position. *American Behavioral Scientist*, *66*(3), 274–291. https://doi.org/10.1177/0002764221996775

Mutz, D. C. (2015). *In-your-face politics: The consequences of uncivil media*. Princeton University Press.

Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, *99*(1), 1–15. https://doi.org/10.1017/S0003055405051452

Nai, A. (2018). Going negative, worldwide: Towards a general understanding of determinants and targets of negative campaigning. *Government & Opposition*, *55*(3), 430–455. https://doi.org/10.1017/gov.2018.32

Nai, A., Medeiros, M., Maier, M., & Maier, J. (2022). Euroscepticism and the use of negative, uncivil and emotional campaigns in the 2019 European Parliament election: A winning combination. *European Union Politics*, *23*(1), 21–42. https://doi.org/10.1177/14651165211035675

Ohr, D., & Oscarsson, H. (2013). Leader traits, leader image, and vote choice. In K. Aarts, A. Blais, & H. Schmitt (Eds.), *Political leaders and democratic elections* (pp. 187–214). Oxford University Press. https://doi.org/10.1093/ACPROF:OSOBL/9780199259007. 003.0011

Otto, L. P., Lecheler, S., & Schuck, A. R. T. (2020). Is context the key? The (non-)differential effects of mediated incivility in three European countries. *Political Communication*, *37*(1), 88–107. https://doi.org/10.1080/10584609.2019.1663324

Pancer, S. M., Brown, S. D., & Barr, C. W. (1999). Forming impressions of political leaders: A cross-national comparison. *Political Psychology*, *20*(2), 345–368. https://doi.org/10.1111/0162–895X.00148

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, *6*(2), 259–283. https://doi.org/10.1177/1461444804041444

Reese, G., Steffens, M. C., & Jonas, K. J. (2013). When black sheep make us think: Information processing and devaluation of in- and outgroup norm deviants. *Social Cognition*, *31*(4), 482–503. https://doi.org/10.1521/SOCO_2012_1005

Rooduijn, M., van Kessel, S., Froio, C., Pirro, A., de Lange, S., Halikiopoulou, D., Lewis, P., Mudde, C., & Taggart, P. (2019). *The PopuList*. https://popu-list.org

Roseman, I. J., Mattes, K., & Redlawsk, D. P. (2021). How the emotion of contempt can help explain political effects of incivility. In A. S. Walter (Ed.), *Political incivility in the parliamentary, electoral and media arena* (pp. 107–123). Routledge.

Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*, *49*(3), 399–425. https://doi.org/10.1177/0093650220921314

Sobieraj, S., & Berry, J. M. (2011). From incivility to outrage: Political discourse in blogs, talk radio, and cable news. *Political Communication*, *28*(1), 19–41. https://doi.org/10.1080/10584609.2010.542360

Stephens, D., Banducci, S., Horvath, L., & Krouwel, A. (2019). The "coarsening" of campaigns. In D. Jackson, E. Thorsen, D. Lilleker, & N. Weidhase (Eds.), *UK election analysis 2019: Media, voters and the campaign* (p. 15). Centre for Comparative Politics and Media Research. https://www.electionanalysis.uk

Stiers, D., Larner, J., Kenny, J., Breitenstein, S., Vallée-Dubois, F., & Lewis-Beck, M. (2021). Candidate authenticity: "To thine own self be true." *Polit-*

ical Behavior, *43*(3), 1181–1204. https://doi.org/10.1007/S11109–019-09589-Y/TABLES/5

Sydnor, E. (2019). Signaling incivility: The role of speaker, substance, and tone. In R. G. Boatright, T. J. Shaffer, S. Sobieraj, & D. Goldthwaite Young (Eds.), *A crisis of civility? Political discourse and its discontents* (pp. 61–80). Routledge.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540–5907.2006.00214.x

The Federal Returning Officer. (2021). *Bundestag election 2021*. https://www.bundeswahlleiter.de/en/bundestagswahlen/2021/ergebnisse.html

van der Eijk, C., van der Brug, W., Kroh, M., & Franklin, M. (2006). Rethinking the dependent variable in voting behavior: On the measurement and analysis of electoral utilities. *Electoral Studies*, *25*(3), 424–447. https://doi.org/10.1016/j.electstud.2005.06.012

Walter, A. (2021). Introduction: The study of political incivility and its challenges. In A. S. Walter (Ed.), *Political incivility in the parliamentary, electoral and media arena* (pp. 1–16). Routledge.

About the Author

Chiara Vargiu is a PhD candidate in political science at the University of Lausanne. Her PhD deals with perceptions and effects of elite incivility, and it is part of the project The rise of Negative Voting, funded by the Swiss National Science Foundation (SNSF). Currently, she is conducting a one-year research stay at the Amsterdam School of Communication Research (ASCoR), University of Amsterdam.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 275–285
https://doi.org/10.17645/pag.v10i4.5701

Article

The Decision to Go Negative: Election Types, Candidate Characteristics, and Electoral Competition

Huang-Ting Yan

Institute of Political Science, Academia Sinica, Taiwan; huangting@gate.sinica.edu.tw

Submitted: 29 April 2022 | Accepted: 6 September 2022 | Published: 30 December 2022

Abstract
This study examined the conditions that motivate candidates to go negative during a parliamentary election campaign. We argue that by-elections encourage candidates to engage in more negative campaigning. Three mechanisms might explain the alleged link: time pressure, media exposure, and voter turnout. Two main factors jointly determine which candidates rely heavily on negative campaigning during by-elections: candidate characteristics and electoral competition. New data collected from press coverage of Taiwanese legislative elections (2008–2022), covering 318 campaigns in single-member electoral districts, were analysed using the qualitative comparative analysis method. We modelled negative campaigning as a combination of a list of potential causal conditions. Thereafter, process-tracing methods were applied to analyse a typical case to demonstrate the internal causal mechanism. The qualitative comparative analysis results and the case study indicate that increased electoral competition causes parachute candidates to criticise political opponents during a by-election campaign, with less emphasis on their own policy proposals. These results suggest that researchers should pay close attention to important contextual factors that underlie candidates' strategic choices, particularly during by-elections.

Keywords
by-elections; candidate characteristics; competition; legislative; negative campaigning; qualitative comparative analysis; Taiwan

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Negative campaigning is defined as any criticism levelled by one candidate against another during a campaign (Geer, 2006, p. 23), in contrast to the use of messages intended to promote one's own policy positions and record. Despite considerable variability between countries, the phenomenon of negative campaigning is observed worldwide (Valli & Nai, 2020). These trends have stimulated political science efforts to understand the impact of attacks, as well as who uses them and under what conditions.

Recent research has identified factors associated with the use of negative campaigning, including micro-level factors, that is, gender (e.g., Herrnson & Lucas, 2006; Walter, 2013), personality traits (e.g., Nai, 2019; Nai & Maier, 2020), party affiliation and ideology (e.g., Elmelund-Præstekær, 2010; Nai & Sciarini, 2018), and incumbency status (e.g., Nai, 2020; Valli & Nai, 2020); and macro-level variables, that is, electoral competitiveness (Fowler et al., 2016; Nai & Sciarini, 2018), time pressure (Nai & Martinez i Coma, 2019; Nai & Sciarini, 2018), and campaign resources (Grossmann, 2009). Based on a recent study to understand the multiple levels of influence on candidates and their use of negativity during election campaigns, the political profile of candidates was the strongest predictor, followed by personality traits and perceived electoral competitiveness (Maier &

Nai, 2021). In general, members of governing parties (Maier & Nai, 2021) and incumbents (Nai, 2020) are less likely to attack, whereas candidates tend to attack ideologically distant rivals (Maier & Nai, 2021; Nai, 2020).

However, these studies had certain limitations. The use of campaign negativity depends not only on microsystems—candidate characteristics—or political, environmental factors. Other broader political conditions influence the structure and availability of microsystems and the manner in which they affect a candidate's decision to "go negative." Although recent research assesses how the context drives or moderates the influence of individual characteristics on candidates' use of negative campaigning (Nai, 2020), we know little about combinations of multiple factors in understanding what influences a candidate's decision to attack their political opponent. Furthermore, we should distinguish the effects of general elections and by-elections—the latter create distinct institutional environments where factors supposed to drive the use of negativity in general election campaigns may exert divergent effects on candidates' strategic considerations. Finally, existing studies focus on different sets of characteristics of the candidates, such as incumbency status, gender, and ideology; a vital feature—that is, candidates parachuted into a constituency—should be considered. Candidates' local ties are expected to drive their strategic calculations.

In this study, we identified possible combinations of causally relevant conditions that drive candidates to "go negative"; to the best of our knowledge, there is no study on this topic. We argue that by-elections encourage candidates to engage in more negative campaigning. Three mechanisms might explain the alleged link: *time pressure*, *media exposure*, and *voter turnout*. Tremendous time pressure and comprehensive media coverage cause candidates to use negative campaigning as a means to increase turnout among their base and attract undecided voters to win the by-election. Then the two main factors—candidate characteristics and electoral competition—jointly influence how the three mechanisms work and set up incentives to attack accordingly. We then modelled negative campaigning as a combination of a list of potential causal conditions, using new data collected from press coverage of Taiwanese legislative elections (2008–2022) and employing the qualitative comparative analysis (QCA) method and a case study. Analyses reveal that higher levels of electoral competition influence parachute candidates' decision to "go negative" during a by-election campaign.

This work builds upon several prior related studies in comparative politics. First, this study identifies possible combinations of conditions on the candidate and context level to understand what influences a candidate's decision to "go negative," thus contributing to existing research on the link between both levels and the use of negative campaigning (Nai, 2020; Valli & Nai, 2020). Second, this study contributes to a broader understanding of the strategic considerations of candidates in different types of elections. Finally, little is known about the consequences of parachuting candidates into party nominations, except for addressing their legislative behaviours (Koop & Bittner, 2011; Russo, 2012). This study is among the first to examine their effects on the use of negativity in election campaigns.

2. Theoretical Arguments

First, this study argues that for a better understanding of the occurrence of negative campaigning, we should focus on the combination of election types and two main factors: candidate characteristics and electoral competition. We differentiate between general elections and by-elections, arguing that political actors are more likely to decide to "go negative" during by-elections compared to general elections. Then, candidate characteristics and electoral competition jointly determine which candidates rely heavily on negative campaigning in by-elections.

Three mechanisms might explain the alleged link: time pressure, media exposure, and voter turnout. First, as the election draws near, political actors are more likely to "go negative" (Nai & Martinez i Coma, 2019; Nai & Sciarini, 2018). At the outset of a campaign, candidates tend to establish their personal brands by providing voters with information about who they are and proposing policies to respond to concerns that are salient for local residents. Then, attacks come towards the end of the campaign in order to draw a sharp contrast (Damore, 2002; Freedman & Goldstein, 2002; Ridout & Holland, 2010). Unlike general elections, where candidates may prepare for the election in advance and devote substantial time to constituency service, by-elections can create significant pressure because of the shorter time to election day. Given the time constraints, candidates are likely to resort to negative campaigning.

Second, negative advertisements make for particularly juicy morsels for the media. In the world of heightened media competition to capture the attention of the "news grazers," election news can be reduced to a conflict between candidates and between journalists and candidates who rail against false charges (Ansolabehere & Iyengar, 1995, p. 134), which appears to be a frequently used tactic to win ratings battles (Vliegenthart et al., 2011, p. 96). Furthermore, the media exaggerates its portrayal of the negativity of political campaigns (Geer, 2012; Hansen & Pedersen, 2008; Ridout & Smith, 2008) by providing extensive coverage of a few negative campaign messages, and, as a result, the general public is presented with a biased view of the political campaign (Hansen & Pedersen, 2008). Therefore, candidates may attract media attention and affect subsequent public opinion dynamics by engaging in negative campaigning. As by-elections are held for fewer seats compared to general elections, we can expect higher levels of negativity as media attention for each constituency increases during by-elections.

Third, competitors engage in negative campaigning to diminish positive feelings for their rivals and increase public favour for themselves (Nai & Walter, 2015; Skaperdas & Grofman, 1995; Walter, 2014). In addition, parties also resort to negative campaigning to mobilise core voters, as this can activate and reinforce party preferences and increase turnout among their base (Stuckelberger, 2021). By attacking another party, a party defines a conflict line that allows its core voters to identify themselves as part of an "us" versus "them" battle, thus creating a stronger party identification among its voters (Elmelund-Præstekær, 2010); as a consequence of the conflict line, parties can draw voters' attention and motivate them to vote (Lau et al., 2007). The effect is much more prevalent for by-elections, where turnout would generally be lower compared to general elections (Gallagher, 1996). This implies that candidates are likely to resort to negative campaigning to boost their base's voter turnout and attract such voters who dislike a candidate (Klein & Ahluwalia, 2005) or are particularly affected by a negative tone.

The presence of by-elections is not sufficient for the negative campaigning outcome. Another determinant is candidate characteristics. A parachute candidate has fewer roots in a given constituency he/she is running to represent and is more likely to "go negative" at by-elections compared to a local candidate who has spent years working for the constituency. Parachute candidates are under tremendous time pressure during by-elections, as they must quickly build connections to the constituency by advertising themselves as the person who identifies more with the community, is more attuned to their concerns and interests, and will promote them. Furthermore, they run the risk of preparing policy proposals that may provoke a backlash from the resident population, partly due to unfamiliarity with local political issues. In this sense, attacks should be more strident as huge media exposure at by-elections makes it more likely for negative advertising to reach a much larger electorate, and low voter turnout necessitates negative advertising to mobilise parties' base. Local candidates, by contrast, tend to emphasise their brokerage services to constituents and policies to respond to concerns most salient for local residents. Moreover, negative campaigning strategies will damage the positive brand image and reputation that they have built over many years.

Another determinant is electoral competition. Candidates with a small number of reliable votes are less motivated to attack their opponents, as attacks do not necessarily turn the tide and can backfire, hurting their own popularity (Kahn & Kenney, 2004; Lau et al., 2007). Specifically, if candidates decide to spend considerable time on constituency management for the next several years, attacking their opponents would damage their self-image. Similarly, candidates with considerable support, which is enough to assure victory, are not tempted to "go negative" as it may alienate potential voters. The negativity of the campaign increases with the competitiveness of the race (Fowler et al., 2016), and we might expect a stronger effect for by-elections, especially when the candidates' base is identical in size. This encourages negative campaigning to boost their respective base's voter turnout and attract those who are specifically impacted by the negativity effect. It could be a key to electoral success under significantly lower turnout at by-elections. In sum, we argue that higher levels of electoral competition inform parachute candidates' decisions to level criticism against competitors during a by-election campaign.

In summary, due to the unfamiliarity with the constituency they are contesting, parachute candidates are less likely to emphasise policy proposals. Moreover, the by-election creates significant pressure because of the shorter time to election day. The by-election, however, offers these candidates the opportunity to attract considerable media attention. Furthermore, by-elections are characterised by low voter turnout, which increases the effectiveness of the strategy to secure the base. The effects, however, would be widespread if there were a high level of electoral competition. Therefore, we hypothesise that the presence of by-elections, combined with the presence of parachute candidates and with higher levels of electoral competition, leads to negative campaigning outcomes. In other words, *if* parachute candidates during a by-election campaign run to represent a given constituency with higher levels of electoral competition, *then* they decide to level criticism against competitors with less emphasis on their own policy proposals.

3. Data and Methods

New data collected from press coverage of Taiwanese legislative elections (2008–2022), covering 318 campaigns in single-member electoral districts, were analysed using the QCA method. The unit of analysis was the candidates in the Taiwanese legislative election. Taiwan's political structure is divided into two camps: the pan-green coalition, led by the Democratic Progressive Party (DPP), and the pan-blue coalition, led by the Kuomintang (KMT). This bipolar competition contributes to the strong trend towards one candidate per camp per district. All candidates running for the two camps, in which the overwhelming majority of candidates belong to the two major political parties, the DPP and the KMT, were analysed. There should have been a total of 636 candidates in the study. However, some candidates were running to represent a given constituency in several legislative elections; a list of 370 candidates was examined accordingly. We sourced the data on the candidate list from the Central Election Commission.

The QCA method was used in this study. It is a configurational method to explore the relationships between the combinations of "conditions" and "outcomes" (Ragin, 2008). Of three types of QCA—crisp-set QCA (csQCA), multi-value QCA (mvQCA), and fuzzy-set

QCA (fsQCA)—this study employs csQCA because all the conditions and outcome factors are bivalent, and their values should generate crisp sets (Thiem, 2014, p. 492). The csQCA method hinges on the logic of Boolean algebra. We must construct a dichotomous data table consisting of 0 or 1 attributed to conditions and outcomes. Then we develop a truth table in which a given outcome corresponds with the combination of conditions. After addressing possible contradictory configurations, the complex, intermediate to parsimonious formula can be produced by Boolean minimisation. Finally, we must interpret the three solution types.

This study included 636 cases—a negative campaign by a candidate (or a non-negative campaign) was considered a case. The outcome factor, NEGCAMP, captures whether candidates use negative campaigning. Candidates debate and defend their own positions and policies and most likely attack their opponents' programs and policy propositions on the hustings. The use of such attacks as a source of coding decisions is inappropriate. Candidates were identified as using negative campaigning if they criticised one of the following aspects of the target: record and accomplishments, controversial issues specific to the campaign, personal characteristics, and physical and socio-demographic attributes. We selected Taiwan's four main newspapers (online versions)—*Apple Daily*, *China Times*, *Liberty Times*, and *United Daily News*—which provided rich election coverage for the data collection. *China Times* is seen as the pan-blue media that highly publicises a smear campaign or a low blow made by the pan-green candidate against his pan-blue opponent, and vice versa for *Liberty Time*, which represents pan-green viewpoints. It is vital to include both newspapers to prevent selection bias. We used a combination of candidate names and negative campaigning-related search terms (e.g., attacks, smear, spread rumours, propaganda) and searched the online version of the four main newspapers. We also searched Google to prevent omitted reports and subsequently found very similar results in other electronic media, such as *ETtoday*, *Newtalk*, and *TVBS* (for a case, see Appendix A in the Supplementary File).

Record and accomplishments involve the candidate's past performance in terms of facilitating local economic development and increasing budgets for road construction and other local projects. For example, Tsai Shih-ying (蔡適應), who ran for a seat in Keelung City Constituency, attacked his opponent Hau Lung-pin (郝龍斌)—former Taipei mayor—saying that the construction bidding for the Taipei Twin Towers failed and the project was repeatedly delayed, and he negatively advertised Hau's second-to-last citizen satisfaction with local government performance; Chuang Suo-hang (莊碩漢), who ran for a seat in New Taipei City Constituency X, presented his competitor, Lu Chia-chen (盧嘉辰)—current legislator—as a legislator who performed poorly, based on the report of legislator evaluation conducted by the Citizen Congress Watch, and a liar who exaggerated his efforts to obtain over 170 billion NT dollars in local construction budget; Lee Chin-yung (李進勇), who ran for a seat in Yunlin County Constituency I, publicly blasted his rival Chang Chia-chun (張嘉郡) for doing nothing to increase the old-age farmer allowance and Yunlin county exclusive fishing right. Chang Chih-ming (張志明), who ran for a seat in Taitung County Constituency, roasted his competitor, Liu Chao-hao (劉櫂豪), saying that as a Taitung legislator for eight years, he was unable to tackle problems such as the construction of Provincial Highway 26, facilitating train ticket purchases for Taitung residents, water supply pervasion, the import of betel pepper, and the sale of public farmland.

Furthermore, record and accomplishments reflect the negative side of a candidate performance. There are some cases where candidates were attacked by their rivals because of potential scams and scandals. These included the Central Motion Picture Corporation case used by Hsu Kuo-yung (徐國勇) as propaganda against Tsai Cheng-yuan (蔡正元), who was suspected of embezzlement; illegal loans from public banks backed by cut-price land for which Lin Yi-shih (林益世) was a suspect and under widespread attacks by Chiu Chih-wei (邱志偉); and a sex scandal about which Ho Po-wen (何博文) made a big story to stigmatise Wu Yu-sheng (吳育昇).

Some controversial affairs arising from the election campaign may become the target, and whether that certainly happened or is fabricated is less important. Candidates filed bribery allegations against their opponents—defamatory forms of communication that involve reputational damage. For example, Chien Chao-tong (簡肇棟), who ran for a seat in Taichung County Constituency III, and Ho Min-hao (何敏豪), who ran for a seat in Taichung City Constituency III, accused his political opponents of voter buying. Candidates also filed illegal canvass allegations against their rivals. For example, Chang Kuo-hsin (張國鑫), who ran for a seat in Nantou County Constituency I, accused the KMT candidate Ma Wen-chun (馬文君) of more than 10 Japanese people openly backing her, on suspicion of violating the Civil Servants Election and Recall Act. Furthermore, Lin Yu-chang (林右昌), who ran for a seat in Keelung City Constituency, heavily criticised his rival Hsieh Kuo-Liang (謝國樑) for utilising Er Xin High School to mobilise and publicise, in which teachers and staff were politically bullied. Finally, some took aim at campaign staff. The DPP candidate Yao Wen-chih (姚文智) accused the KMT Legislator Huang Chao-shun's (黃昭順) assistant of involvement in the Huaxin Laundry arson case.

Personal characteristics imply non-physical qualities or features of candidates and make them recognisable. For example, Lin Chien-jung (林建榮) and Chen Ou-po (陳歐珀), who ran for a seat in Yilan County Constituency, mutually described each other as "political speculators." Furthermore, Su Chen-ching (蘇震清), who ran for a seat in Pingtung County Constituency I, disseminated information alleging that his competitor, Tsai Hau (蔡豪), was more outrageous than Zheng Taiji—former

Pingtung County Council Speaker closely associated with the criminal underworld. Sun Ta-chien (孫大千), who ran for a seat in Taoyuan City Constituency VI, accused Chao Cheng-yu (趙正宇) of withdrawing from the KMT and running for the election as a betrayer and being disloyal to his party.

Physical attributes refer to someone's physical or facial features, while socio-demographic attributes refer to a combination of social and demographic factors that define people in a specific group or population, including age, gender, ethnicity, education level, income, family ties, location, and so forth. For example, Ting Shou-chung (丁守中), who ran for a seat in Taipei City Constituency I, criticised his opponent Wu Szu-yao (吳思瑤) as "a woman aged 40 to 50 who called herself sister Szu-yao when she met a child." Lin Yu-fang (林郁方), who ran for a seat in Taipei City Constituency V, bombarded Lim Tshiong-tso (林昶佐) for being mentally abnormal and having longer hair than a woman. Lee Yung-ping (李永萍), who ran for a seat in New Taipei City Constituency XII, attacked Lai Pin-yu (賴品妤) for being too young to be a legislator. Hung Chun-yi (洪宗熠), who ran for a seat in Changhua County Constituency III, and his fans club hung a black banner reading "If Cheng Ru-fen (鄭汝芬) (family) does not fail, Changhua will never improve," to highlight the notorious Hsieh family who dominated southern Changhua over three generations.

We collected the data from press coverage instead of candidates' self-reports for the following reasons. First, as Maier and Nai (2021, p. 8) indicated: "Self-report assessments on the use of negative campaigning might suffer from validity issues. Because voters usually report their dislike of attacks, candidates might be motivated to downplay the level of negativity they may have employed in their campaigns." Second, media coverage is significantly higher for candidates who go negative (Maier & Nai, 2020) or rely on personal attacks (Gerstlé & Nai, 2019), and this provides a more reliable source of information for the analysis of the drivers of negative campaigning. Appendix B in the Supplementary File contains the list of negative campaigning and data source.

We have three bivalent conditions, of which BYELEC (by-election campaign) is already of a dichotomous nature, where 1 = *a by-election*, defined as an election held to fill a vacant electorate seat if a member of Parliament (MP) resigns or dies or becomes ineligible to continue in office, and 0 = *otherwise*. PARACHU (parachute candidates) delineates an election candidate who does not reside in and has little connection to the area they are running to represent. This included three types of candidates: a political neophyte, one who served as a member of the national parliament or government but has no connection to the constituency they wish to stand in, and party-list representatives. For example, in the 2012 legislative election, Chien Wei-chuan (錢薇娟)—a female basketball player—ran for a seat in New Taipei City Constituency II, representing the KMT; Su Jun-bin (蘇俊賓)—the director general of the Department of Environmental Protection, Taoyuan County, and the director general of the Government Information Office, Executive Yuan—ran for a seat in Tainan City Constituency IV, representing the KMT; Chiu Yi-ying (邱議瑩)—a member of the national parliament from Pingtung County Constituency—ran for a seat in Kaohsiung City Constituency I, representing the DPP; and Cheng Li-wun (鄭麗文)—a party-list legislator—ran for a seat in Taichung City Constituency VII, representing the KMT. These candidates are considered to be parachuted into the MP post. The data on candidates' objective backgrounds were collected from election bulletins. Appendix C in the Supplementary File contains the list of parachute candidates.

ELECOMP (electoral competition) records the level of electoral competition. The study calculated the level of political competition in electoral regions, comparing the percentage of votes received by candidates formally nominated from the two major camps per legislative election between 2008 and 2022 per district. A narrow margin between two candidates shows a high level of political competition. We did not use the gap in the previous election as a measure of the level of electoral competition, as it cannot reflect the current electoral competitiveness. The level of electoral competition is determined by actual electoral data as a proxy for pre-election closeness. It is likely that negative campaigning influences actual votes. The level of electoral competition—the base variable is continuous—could be calibrated into a dichotomous variable to reduce the endogenous effect. There are, however, a few un-nominated aspirants. The study identified candidates who participated in elections without the DPP or KMT approval and merged their votes into the base of the two camps according to their political spectrum. ELECOMP is coded 1 when the difference between two candidates in the percentage of votes is less than 5%. The data were collected from the Central Election Commission.

The truth table shows that some configurations are associated with a contradictory outcome in which cases are coded 1 on the outcome of interest, and others are 0. There are three ways to handle this issue. It can be resolved using statistical methods that compare the distribution of outcomes for a given configuration and that of the outcomes for cases not captured by that configuration (Roscigno & Hodson, 2004, p. 25). Alternatively, it can also be achieved by assigning an outcome to the contradictory configuration according to the outcome value shown by the majority of cases (Skaaning, 2011, p. 402). A third option is the choice of a consistency threshold that separates the sufficient truth table rows from those designated as insufficient for the outcome. We applied a consistency criterion of 0.6. The three strategies were adopted in the study. Some logically possible configurations exist in which non-observable cases fit, and these are defined as logical remainders. Conventionally, this problem can be tackled using three

strategies. Of these three, a conservative approach is to treat all logical remainders as *false*. Meanwhile, another strategy involves treating these remainders as *do not care*, so as to generate the most parsimonious formula. Lastly, there is the third approach to finding a solution term of intermediate complexity, which is based on existing knowledge supporting the idea that the presence of conditions is linked to outcome 0 or 1 (Ragin & Sonnett, 2004). It is necessary to produce and report three formulae: complex, intermediate, and the most parsimonious, according to standards of good practice in QCA (Schneider & Wagemann, 2010).

To check robustness, one can use reduction, addition, or replacement of explanatory factors to test if similar results would emerge (Skaaning, 2011). Herrnson and Lucas (2006) show that male candidates are more likely to attack their political opponent(s) compared to female candidates, but Walter (2013) finds no difference. This study, therefore, added the other causal condition, GENDER, where 1 = *male candidates* and 0 = *otherwise*, which is supposed to drive the use of negative campaigning. Furthermore, it was found that negative campaigning has increased over time in the US (Fowler et al., 2016, p. 53), while the trend was less clear-cut for Europe (Walter, 2014). Therefore, this study used the other causal condition, YEAR, where 1 stands for the *years after 2020* and 0 *otherwise*, which is likely to influence the use of negative campaigning strategies. Moreover, there is evidence pointing out that challengers are less likely to run positive campaigns than incumbents due to a reduced likelihood of promoting their record and accomplishments (Nai, 2020; Nai & Walter, 2015). We, however, did not include this causal condition, given that all candidates were challengers at by-elections. Alternatively, robustness can also be checked by altering the calibration thresholds (Glaesser & Cooper, 2014; Skaaning, 2011). ELECOMP can be transformed into a condition consisting of multiple categories where we assign a value of 2 to cases if the difference between the two candidates in the percentage of votes is less than 5%, a value of 1 with a 5% to 10% difference, and a value of 0 with the difference of more than 10%. To deal with multi-value conditions, this study employs the mvQCA. Lastly, one can check robustness by altering the frequency thresholds (Skaaning, 2011). An analysis can be run where the frequency thresholds are raised to two cases, compared to the original analysis that used all configurations representing at least one case.

The next step of this study involved identifying cases to discern causal mechanisms. After reporting results obtained using csQCA, we followed Schneider and Rohlfing's (2013, 2016) proposal for selecting appropriate cases for in-depth studies. At least one case has to be chosen for each term of solution (the principle of diverse case selection). In terms of typical case selection, this study focused on cases that are members of just one term of the solution (the principle of unique membership; Schneider & Rohlfing, 2013, 2016).

4. Results

As shown in Table 1, when considering three conditions influencing candidates' use of negative campaigning, the analysis found that the combination of three conditions accounts for the use of negative campaigning (BYELEC * PARACHU * ELECOMP → NEGCAMP). Therefore, during a by-election campaign (BYELEC), higher levels of electoral competition (ELECOMP) lead to parachute candidates' (PARACHU) decisions to level criticism against competitors (NEGCAMP), with less emphasis on their own policy proposals.

There are other causal combinations of conditions that can account for the use of negative campaigning (e.g., byelec * parachu * ELECOMP * GENDER → NEGCAMP). However, these were not identified in all solution types. The same procedure was used to obtain the formula for [0: a non-negative campaign] configurations. We found at least one of the three conditions, that is, byelec, parachu, and elecomp, in large proportion of the complex, intermediate, and parsimonious formula. Compared to the path BYELEC * PARACHU * ELECOMP covering the observed [1: a negative campaign] outcome cases, we found that this type of path was not associated with the [0: a non-negative campaign] outcome (Appendix F in the Supplementary File).

Next, we chose appropriate cases for analysis of causal mechanisms. This study applied the principle of diverse case selection and the principle of unique membership for the choice of typical cases—that is, "Taichung City Constituency II, 2022 legislative by-election"—as the causal process of the term BYELEC * PARACHU * ELECOMP. In the case, we also presented that (a) parachute candidates were less likely to emphasise policy proposals, (b) the by-election created significant time pressure for parachute candidates, and (c) the by-election offered parachute candidates the opportunity to attract considerable media attention and use the strategy of securing the base during close races.

The 2022 legislative by-election was held on 9 January 2022. Its purpose was to elect one member of the Legislative Yuan who would serve the remaining term until 2024, as Chen Po-wei (陳柏惟), who in the 2020 legislative elections defeated the KMT incumbent Yen Kuan-heng (顏寬恒) in the Taichung City Constituency II, was recalled on 23 October 2021. The KMT fielded Yen Kuan-heng to contest the seat again, while the DPP put forward former party-list legislator Lin Ching-yi (林靜儀). Regarding the political careers of the two, in 2001, Yen helped run the first legislative campaign of his father, Yen Ching-piao (顏清標), and also worked as the elder Yen's legislative assistant. Yen Ching-piao was expelled from the Legislative Yuan as a result of his sentence, necessitating a by-election on 26 January 2013 that was won by Yen Kuan-heng. The Yens are the most famous factional family dominating Taichung's coastline and ruling the city district for almost two decades. KMT provided the Yens subsidies in the gravel and casino businesses while also

Table 1. The decision to go negative: QCA.

	Resolving Contradictory Configurations		
	Assigning an outcome to the contradictory configuration according to the outcome value shown by the majority of cases	A consistency threshold of 0.6 that separates the sufficient truth table rows from those designated as insufficient for the outcome	Ratios of configuration to non-configuration percentages
Original			
Condition: BYELEC, PARACHU, ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP
Robustness			
Addition of condition: GENDER	C[1]: BYELEC * PARACHU * ELECOMP + byelec * parachu * ELECOMP * GENDER P[1]: BYELEC * PARACHU * ELECOMP + byelec * parachu * ELECOMP * GENDER	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP + byelec * parachu * ELECOMP * GENDER P[1]: BYELEC * PARACHU * ELECOMP + byelec * parachu * ELECOMP * GENDER
Addition of condition: YEAR	C[1]: BYELEC * PARACHU * ELECOMP I[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP I[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP I[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP
Change the Calibration Thresholds: ELECOMP	C[1]: BYELEC{1} * PARACHU{1} * ELECOMP{2} + BYELEC{0} * PARACHU{0} * ELECOMP{1} P[1]: BYELEC{1} * PARACHU{1} * ELECOMP{2} + BYELEC{0} * PARACHU{0} * ELECOMP{1}	C[1]: BYELEC{1} * PARACHU{1} * ELECOMP{2} P[1]: BYELEC{1} * PARACHU{1} * ELECOMP{2}	C[1]: BYELEC{1} * PARACHU{1} * ELECOMP{2} + BYELEC{0} * PARACHU{0} * ELECOMP{1} P[1]: BYELEC{1} * PARACHU{1} * ELECOMP{2} + BYELEC{0} * PARACHU{0} * ELECOMP{1}
Change the Frequency Thresholds: Two Cases	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP	C[1]: BYELEC * PARACHU * ELECOMP P[1]: BYELEC * PARACHU * ELECOMP

Notes: C[1]—complex formula for the positive [1] outcome, I—intermediate formula, P—parsimonious formula; to find the intermediate formula, this study applied a more rational assumption that the presence of the years after 2020, as "easy" counterfactual, is associated with candidates' decisions to go negative (YEAR{1} → O{1}); for csQCA, an uppercase letter represents that the condition is present, whereas a lowercase letter indicates the absence; for mvQCA, conditions are expressed in capital letters, their value is presented next to them in brackets; please see Appendix F in the Supplementary File for all analyses and the solutions for the negative outcomes [0].

offering them the opportunity to serve in public offices in exchange for their support to ensure the continued dominance of KMT. His sister, Yen Li-min (顏莉敏), is the current vice-speaker of the Taichung City Council.

Lin—a physician-turned-politician—has a reputation for being clean and idealistic, having previously been a party-list legislator and the head of several departments within the DPP, including departments of gender equality and international affairs. Lin, however, has little connection to the Taichung City Constituency II. The fact of being parachuted into the constituency made it less likely for her to claim a record and emphasise her policy proposals. It was not until 21 days before election day that Lin shared her vision for the people of the Taichung City Constituency II and outlined her policy priorities for the next four years. According to Yen, her competitor: "Over the last month or so, Lin did not put forward visions, goals, and long-term plans dealing with voters' needs. It was too late to propose it now" (S1: Data Source 1, Appendix E in the Supplementary File). Although Lin had posted a policy outline—the so-called 10 good recipes (十帖良方)—on Facebook (S2), she also admitted the small number of views, clicks, comments, likes, and shares on this post, implying people's disinterest in her policy outline (S3). Furthermore, her policy propositions deviated from the party script. She said: "After being elected, I am definitely sure that I am going to arrange the oral cancer screening for residents in Taichung's coastline. Betel nut chewing is seen everywhere that really broke me down" (S4). This criticism of the local residents sparked a harsh public reaction.

The by-election, however, offered Lin the opportunity to use the resources of the state and the ruling party to attack and smear Yen and his family. First, the by-election creates significant pressure because of a shorter time to election day, which contributes to the DPP's all-out mobilisation efforts in support of Lin's campaign (S5). Former Taichung mayor Lin Chia-lung (林佳龍) served promptly as chairman of the campaign office (S6). Ker Chien-ming (柯建銘), the majority leader of the Legislative Yuan, mobilised DPP legislators to campaign for Lin (S7), while President Tsai Ing-wen (蔡英文) and Vice President Lai Ching-te (賴清德) stumped for Lin ahead of elections several times (S8–10). It was a team that combined mobilisation, propaganda, and information warfare to launch an all-out attack against the Yen. As Lin indicated in her victory celebration: "The election campaign period is extremely short, and it is a very difficult constituency to contest. This victory represents the best unity of the DPP" (S11).

Second, it is the only national election and attracts considerable media attention. Given the heightened media competition to capture the attention of "news grazers," the focus remains on the negative aspects of the campaign: the attacks and dramatics of the opposing sides. By publicly criticising the Yen family through media addresses and public speeches, Lin depicted the family as a family of gangsters standing at the intersection of organised crime, electoral politics, and business tycoons (S12), which will affect subsequent public opinion dynamics. During the by-election, various political talk shows discussed Yen's background as the scion of a family of gangsters and his family's political connections; media coverage of Yen was generally unfavourable (S13–14). At a press conference on 13 December 2021, Yen released statistical data on the number of days he was attacked by shows, for example, Taiwan Front Line (台灣最前線, 37 days), Coco Hot News (辣新聞152, 30 days), and Taiwan Go for It (台灣向前行, 25 days; S15). Furthermore, for decades, the Yen family—through its connections and control over vote captains, loyalty networks, and information flows—has had considerable sway in Taichung and could always be relied upon to leverage its influence in the locality. However, considerable media coverage during the by-election caused the Yen family's control over information flows to slip, motivating Lin to attack them publicly, which contributed to the local electorate's disillusionment with the Yens (S16).

Third, by-elections are characterised by low voter turnout, which increases the effectiveness of the strategy of securing the base. The effects, however, would be widespread if there is a high level of electoral competition. The turnout in the 2013 legislative by-election held for Taichung City Constituency II was 48.89%, compared to the general elections in 2016 (70.80%) and 2020 (77.20%). Yen won by a narrow margin of 1% in 2013 and 3% in 2016, whereas Chen won a narrow victory by a 3% margin in 2020. In Taichung City Constituency II, which comprises Dadu (大肚), Longjing (龍井), Shalu (沙鹿), Wufeng (霧峰), and Wuri (烏日), the DPP advantage in Wuri is overwhelming; Dadu, Longjing, and Shalu are Yen family traditional strongholds; and Wufeng is a swing district, shifting from slightly blue to slightly green in recent years. Given the identical base size for both camps, the DPP launched an all-out attack against the Yen family from the beginning of the election campaign. There were allegations that the Yens' mansion was built illegally on national property (S17), that an MRT station was proposed near a property owned by the Yen family (S18–19), that the Dajia Jenn Lann Matsu Temple (大甲鎮瀾宮) run by the family had been filing the exact same tax returns for decades (S20), etc. This was done to boost their base's voter turnout and create an environment of hatred among young voters towards the Yen family (S21). It eventually stimulated young voters who dislike the KMT's black gold politics—which, in Taiwan, refers to the obtaining of money (the "gold") through an illegal method (hence the gold being "black")—to return home and vote.

Yen—a native of Shalu—emphasised the history of long-term local service and policies to respond to the most salient concerns of local residents to attract local sympathy votes. The Yen camp's strategy was to rely on traditional supporters' votes by exhorting them to vote (S21). In sum, the case study supports the contention that during a by-election campaign, increased

electoral competition causes parachute candidates to criticise political opponents.

5. Conclusions

What drives the use of negative campaigning? Using new data collected from press coverage of Taiwanese legislative elections (2008–2022)—and combining the QCA method and a case study—we find that higher levels of electoral competition cause parachute candidates to criticise political opponents during by-election campaigns.

The results suggest a more comprehensive approach to the study of the drivers of negativity. No single factor can determine whether a political actor will "go negative." Our approach is to identify two broader sets of factors—as Maier and Nai (2021) indicated—and list the potential causal conditions that explain the occasions when candidates opt to "go negative." We did not discuss the significant drivers of negative campaigning during general elections, which may have alternative explanations. Parachute candidates do not go significantly more negative than local candidates during general elections, as there is adequate time to self-advertise among voters and inadequate media coverage to launch a negative political campaign. Personality traits and ideology matter; research has demonstrated that negativity is linked to candidates' personalities and ideological positions (Elmelund-Præstekær, 2010; Nai, 2019; Nai & Maier, 2020; Nai & Sciarini, 2018). Detailed data on candidates' social profiles may enable the examination of causal combinations that can account for the decision to go negative in general elections.

These results also suggest that researchers should pay close attention to important contextual factors that underlie candidates' strategic choices, particularly during by-elections. At least two policy implications warrant consideration. First, research literature indicates that negative campaigning has detrimental effects on the overall political system as it tends to reduce citizens' sense of political efficacy and trust in government and adversely impacts the overall public mood (Lau et al., 2007). Thus, parties need to consider their decision to field parachute candidates during by-elections, as this is more likely to lead to attacks on political rivals. Second, as a consequence of the first rationale, the government should take appropriate actions to raise one's sense of political efficacy and restore public confidence in the government in an increasingly competitive political environment where parachute candidates contest by-elections.

This study has several limitations. First, it only considered the case of Taiwan, which may limit the results' generalisability. Second, we confirmed the arguments using objective data from press coverage. However, as Maier and Nai (2021, p. 7) indicated: "Candidate's perception might affect how he campaigns—even if his perception does not match reality." Future research should test the validity of the proposed arguments using subjective data from a comprehensive survey of candidates. Third, we did not consider a dynamic pattern of negative campaigning—candidates' campaigning against each other, which causes a certain dynamic—or control for time pressure based on the findings that negative campaigning increases as the voting day is close (Nai & Martinez i Coma, 2019; Nai & Sciarini, 2018). We will test the robustness when data become available. Fourth, we provided evidence that by-elections were characterised by the shorter time to election day, especially after candidates were nominated, and by a lower turnout (Appendix D in the Supplementary File). However, no data exist for unpacking the mechanisms underlying how the shorter time to election day would give candidates less time to prepare their campaign and get familiar with the constituency, which in turn might influence candidates' decision to go negative. Fifth, more cases were required to back the claim that the media would be more focused on these if there are only a few districts in which an election takes place. Sixth, there is enormous variation across negative campaigning. Some prefer criticising record and accomplishments of other candidates, while others prefer character attacks (Maier & Nai, 2021). Future research considering such issues would enrich our understanding of a combination of causally relevant conditions that generate the outcome.

Acknowledgments

I am grateful to Taiwan's four main newspapers (*Apple Daily*, *China Times*, *Liberty Times*, and *United Daily News*) for providing rich election coverage. I would like to thank Prof. Yu-Shan Wu (Institute of Political Science, Academia Sinica) and anonymous reviewers for their precious inputs. All remaining errors are mine.

Conflict of Interests

The author declares no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Ansolabehere, S., & Iyengar, S. (1995). *Going negative: How political advertisements shrink and polarize the electorate*. The Free Press.

Damore, D. F. (2002). Candidate strategy and the decision to go negative. *Political Research Quarterly*, *55*(3), 669–685.

Elmelund-Præstekær, C. (2010). Beyond American negativity: Toward a general understanding of the determinants of negative campaigning. *European Political Science Review*, *2*(1), 137–156.

Fowler, E. F., Franz, M. M., & Ridout, T. N. (2016). *Political advertising in the United States*. Westview Press.

Freedman, P., & Goldstein, K. (2002). Lessons learned campaign advertising in the 2000 elections. *Political Communication*, *19*(1), 5–28.

Gallagher, M. (1996). By-elections to Dáil Éireann 1923–96: The anomaly that conforms. *Irish Political Studies*, *11*(1), 33–60.

Geer, J. G. (2006). *In defence of negativity: Attack ads in presidential campaigns*. University of Chicago Press.

Geer, J. G. (2012). The news media and the rise of negativity in presidential campaigns. *PS: Political Science & Politics*, *45*(3), 422–427.

Gerstlé, J., & Nai, A. (2019). Negativity, emotionality and populist rhetoric in election campaigns worldwide, and their effects on media attention and electoral success. *European Journal of Communication*, *34*(4), 410–444.

Glaesser, J., & Cooper, B. (2014). Exploring the consequences of a recalibration of causal conditions when assessing sufficiency with fuzzy set QCA. *International Journal of Social Research Methodology*, *17*(4), 387–401.

Grossmann, M. (2009). Going pro? Political campaign consulting and the professional model. *Journal of Political Marketing*, *8*(2), 81–104.

Hansen, K. M., & Pedersen, R. T. (2008). Negative campaigning in a multiparty system. *Scandinavian Political Studies*, *31*(4), 408–427.

Herrnson, P. S., & Lucas, J. C. (2006). The fairer sex? Gender and negative campaigning in U.S. elections. *American Politics Research*, *34*(1), 69–94.

Kahn, K. F., & Kenney, P. J. (2004). When do candidates go negative? In K. F. Kahn & P. J. Kenney (Eds.), *No holds barred: Negative campaigning in U.S. Senate campaigns* (pp. 19–37). Pearson.

Klein, J. G., & Ahluwalia, R. (2005). Negativity in the evaluation of political candidates. *Journal of Marketing*, *69*(1), 131–142.

Koop, R., & Bittner, A. (2011). Parachuted into parliament: Candidate nomination, appointed candidates, and legislative roles in Canada. *Journal of Elections, Public Opinion and Parties*, *21*(4), 431–452.

Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. *Journal of Politics*, *69*(4), 1176–1209.

Maier, J., & Nai, A. (2020). Roaring candidates in the spotlight: Campaign negativity, emotions, and media coverage in 107 national elections. *The International Journal of Press/Politics*, *25*(4), 576–606.

Maier, J., & Nai, A. (2021). Mapping the drivers of negative campaigning: Insights from a candidate survey. *International Political Science Review*. Advance online publication. https://doi.org/10.1177/0192512121994512

Nai, A. (2019). Disagreeable narcissists, extroverted psychopaths, and elections: A new dataset to measure the personality of candidates worldwide. *European Political Science*, *18*(2), 309–334.

Nai, A. (2020). Going negative, worldwide: Towards a general understanding of determinants and targets of negative campaigning. *Government & Opposition*, *55*(3), 430–455.

Nai, A., & Maier, J. (2020). Dark necessities? Candidates' aversive personality traits and negative campaigning in the 2018 American midterms. *Electoral Studies*, *68*, Article 102233.

Nai, A., & Martinez i Coma, F. (2019). Losing in the polls, time pressure, and the decision to go negative in referendum campaigns. *Politics & Governance*, *7*(2), 278–296.

Nai, A., & Sciarini, P. (2018). Why "going negative"? Strategic and situational determinants of personal attacks in Swiss direct democratic votes. *Journal of Political Marketing*, *17*(4), 382–417.

Nai, A., & Walter, A. S. (2015). *New perspectives on negative campaigning: Why attack politics matters*. ECPR Press.

Ragin, C. C. (2008). *Redesigning social inquiry: Set relations in social research*. University of Chicago Press.

Ragin, C. C., & Sonnett, J. (2004). Between complexity and parsimony: Limited diversity, counterfactual cases, and comparative analysis. In S. Kropp & M. Minkenberg (Eds.), *Vergleichen in der Politikwissenschaft* [Comparisons in political science] (pp. 180–197). Springer.

Ridout, T. N., & Holland, J. L. (2010). Candidate strategies in the presidential nomination campaign. *Presidential Studies Quarterly*, *40*(4), 611–630.

Ridout, T. N., & Smith, G. R. (2008). Free advertising: How the media amplify campaign messages. *Political Research Quarterly*, *61*(4), 598–608.

Roscigno, V. J., & Hodson, R. (2004). The organizational and social foundations of worker resistance. *American Sociological Review*, *69*(1), 14–39.

Russo, F. (2012). The constituency as a focus of representation: Studying the Italian case through the analysis of parliamentary questions. In S. Martin & O. Rozenberg (Eds.), *The roles and function of parliamentary questions* (pp. 32–43). Routledge.

Schneider, C. Q., & Rohlfing, I. (2013). Combining QCA and process tracing in set-theoretic multi-method research. *Sociological Methods & Research*, *42*(4), 559–597.

Schneider, C. Q., & Rohlfing, I. (2016). Case studies nested in fuzzy-set QCA on sufficiency formalizing case selection and causal inference. *Sociological Methods & Research*, *45*(3), 526–568.

Schneider, C. Q., & Wagemann, C. (2010). Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets. *Comparative Sociology*, *9*(3), 397–418.

Skaaning, S. E. (2011). Assessing the robustness of crisp-set and fuzzy-set QCA results. *Sociological Methods & Research*, *40*(2), 391–408.

Skaperdas, S., & Grofman, B. (1995). Modeling negative campaigning. *American Political Science Review*, *89*(1), 49–61.

Stuckelberger, S. (2021). Mobilizing and chasing: The voter targeting of negative campaigning—Lessons from the Swiss case. *Party Politics*, *27*(2), 341–350.

Thiem, A. (2014). Navigating the complexities of qualitative comparative analysis: Case numbers, necessity relations and model ambiguities. *Evaluation Review*, *38*(6), 487–513.

Valli, C., & Nai, A. (2020). Attack politics from Albania to Zimbabwe. A large-scale comparative study on the drivers of negative campaigning. *International Political Science Review*. Advance online publication. https://doi.org/10.1177/0192512120946410

Vliegenthart, R., Boomgaarden, H. G., & Boumans, J. W. (2011). Changes in political news coverage: Personalization, conflict and negativity in British and Dutch newspapers. In K. Brants & K. Voltmer (Eds.), *Political communication in postmodern democracy* (pp. 92–110). Palgrave Macmillan. vWalter, A. S. (2013). Women on the battleground: Does gender condition the use of negative campaigning? *Journal of Elections, Public Opinion and Parties*, *23*(2), 154–176.

Walter, A. S. (2014). Choosing the enemy: Attack behaviour in a multiparty system. *Party Politics*, *20*(3), 311–323.

About the Author

Huang-Ting Yan is a postdoctoral scholar at the Institute of Political Science at Academia Sinica. He is a political scientist and a social epidemiologist interested in the link between political institutions, socioeconomic outcomes, and public health. Dr Yan's work has appeared in political science and public health journals, including *Democratization*, *European Journal of Public Health*, *European Political Science Review*, *International Journal of Public Health*, *International Political Science Review*, *Public Health Nutrition*, and more.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 286–298
https://doi.org/10.17645/pag.v10i4.5718

Article

The Role of Gender in Parliamentary Attacks and Incivility

Željko Poljak

Department of Political Science, University of Antwerp, Belgium; zeljko.poljak@uantwerpen.be

Submitted: 29 April 2022 | Accepted: 22 July 2022 | Published: 30 December 2022

Abstract
It has been well established that politicians attack their competitors to reach their political goals. As such, there is a considerable amount of literature on their attack behaviour. However, this literature almost exclusively investigates attack behaviour during campaigns, and so far, few studies have addressed the nature of attacks during more routine times in parliaments. This article aims to fill this gap by examining in-parliament attack behaviour and, more specifically, the gender characteristics of attacks. It is theorised that women are less likely to attack and be attacked than men due to the stereotypical gender roles. However, it is anticipated that this compliance to stereotypes diminishes as proximity to elections increases, resulting in women engaging in attacks as much as men. To limit the cost of their divergence, attacks employed by and toward women are expected to be more civil. Lastly, this study argues that adherence to gender stereotypes is stronger in countries with candidate-centred parliamentary systems than party-centred ones. This study finds support for the theoretical framework using longitudinal data on individual attacks in the parliaments of Belgium, Croatia, and the UK. Results confirm that politicians adhere to gender stereotypical roles in parliaments, with women attacking and being targeted less than men, and when women do attack or are targeted, less incivility is employed. Proximity to elections makes both women and men more hostile, but women lower the cost of their increasing attack behaviour by using less incivility, unlike men who increasingly opt for uncivil attacks closer to elections. Additionally, these findings strongly apply in the candidate-centred system of the UK, whereas in the party-centred system of Belgium and Croatia, hardly any support for the theory can be found.

Keywords
attacks; incivility; gender; parliaments

Issue
This article is part of the issue “Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics” edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Politicians use attacks to discredit their competitors and to move toward their political goals. For example, politicians may attack, hoping to lower competitors’ approval to secure re-election, entry to office, and policy implementation. As such, much has been written regarding individuals that attack during campaigns, especially through the lens of gender. By surveying politicians, studies have shown how men prioritise attacks during campaigns more compared to women (Herrnson & Lucas, 2006; Maier & Nai, 2021). On the other hand, content studies of campaign messages show that women are known to engage in attacks equally (e.g., Auter & Fine, 2016; Banwart & Bystrom, 2022; Maier, 2015; Walter, 2013) or even more than men (e.g., Evans et al., 2014; Wagner et al., 2017). Despite this abundance of studies, we lack knowledge regarding the gender characteristics of attacks outside campaigns.

Only a handful of recent studies have tackled possible attack behaviour outside campaigns. Focusing on parliamentary speeches, these studies highlighted that men use *adversarial* (Hargrave & Langengen, 2021) and *negative* (Haselmayer et al., 2021) speeches more often than women, which is in line with stereotypical gender roles that see men as more aggressive or dominant (Eagly &

Karau, 2002). Although these studies provide a key indication of the gender characteristics of attackers in parliaments, that is, that men probably attack more compared to women, we still do not know who is at the receiving end of these attacks nor how attack behaviour evolves throughout the electoral cycle. Additionally, far too little attention has been paid to how these attacks are executed, especially when attacks diverge from expected gender roles. For example, women in the parliamentary opposition may choose to attack since it is their role to hold the government accountable (De Giorgi & Ilonszki, 2018). Lastly, we still lack a comparative perspective on this subject because previous studies focused their analyses on single-country cases. This limits our knowledge on the subject, given that gender can play a different role across different political systems.

To provide an understanding of these open questions, I follow the role congruency theory of prejudice by Eagly and Karau (2002), which argues that deviations from stereotypical gender roles may cause women to face prejudice. As society considers women as *communal* (e.g., kind) and men as *agentic* (e.g., aggressive), female politicians showing agentic behaviour may end up not reaching their political goals. This is why men are usually considered more likely to attack than women, and this notion appears to hold in parliaments looking at the forms of speeches (Hargrave & Langengen, 2021; Haselmayer et al., 2021). The first aim of this article is to extend this theoretical framework toward targets of attacks. I expect that gender stereotypes also apply to targets, with women receiving fewer attacks than men. Furthermore, I argue that this gender-conforming behaviour loses its importance as proximity to elections increases, with women and men engaging equally in attack behaviour (Maier, 2015; Walter, 2013).

The second aim of this article is to investigate the manner of attacks in cases when women do assume an *agentic* role, both as an attacker and as a target. I expect incivility, which can be present or absent in an attack, to be the key. Women avoid the cost of showing *agentic* behaviour by using less incivility when they attack compared to men. In turn, all politicians avoid the cost of targeting women, perceived as *communal*, by using less incivility. Lastly, I integrate this framework with the literature on the politics of legislative debate (Fernandes et al., 2021), arguing that adherence to stereotypical gender roles is stronger in parliaments oriented at candidates rather than parties.

These hypotheses are tested for the parliaments of Belgium (2010–2020), Croatia (2010–2021), and the UK (2010–2020). I use data on attacks and incivility employed by individual politicians during parliamentary question time sessions (QTSs). Results show that women are indeed less likely to attack or be attacked than their male colleagues. Women are also less likely to use incivility when they attack, and are less likely to be attacked in an uncivil manner when compared to men. Furthermore, both men and women engage in attacks more frequently as elections approach, but women compensate for this by using less incivility, unlike men, who are more likely to employ incivility closer to elections. Lastly, the comparative design of this study confirms that adherence to gender stereotypes is much stronger in the UK, a country with a political system in which candidates independently run for office in single-member districts. In the party-driven systems of Belgium and Croatia, in which citizens vote for parties and not candidates, politicians are less likely to conform to gender stereotypes. As such, these results provide a valuable understanding of the role gender can play in attacks and the incivility used in parliamentary venues.

2. Attack Politics in Parliaments: Gender Perspective

To analyse the role of gender in parliamentary attacks, I rely on the role congruency theory of prejudice by Eagly and Karau (2002). This theory argues that women face prejudice based on (a) how they behave (descriptive prejudice) and (b) how they should behave (prescriptive prejudice). These prejudices are rooted in stereotypical gender roles that see women as *communal* (i.e., kind, sympathetic, friendly, gentle) and men as *agentic* (i.e., aggressive, dominant, self-confident). Therefore, for instance, if a woman diverges from *communal* behaviour toward *agentic* behaviour, this would negatively impact her reputation (Eagly & Karau, 2002, p. 576).

This broad notion was used by scholars who explored gender differences in attacks during campaigns. Through a survey method with politicians, some studies have demonstrated that female candidates are hesitant to employ attacks in their campaigning strategies (Herrnson & Lucas, 2006; Maier & Nai, 2021). However, content studies of campaigns generally show women to be equally negative as men (Bystrom, 2004). For example, a study of the recent 2020 US Senate race has shown that both female and male candidates used an equal number of attacks in TV ads (Banwart & Bystrom, 2022). At the same time, experts rated Trump's and Clinton's campaigns during the 2016 presidential elections as negative (Nai & Maier, 2018). Furthermore, a study on attack behaviour in party broadcasts in the UK, Netherlands, and Germany found no differences between the attacks made by parties with female and male leaders (Walter, 2013); a similar finding can be observed in German television debates (Maier, 2015). Some studies have even shown female politicians to be more likely to attack than men (e.g., Evans et al., 2014; Wagner et al., 2017). These non-stereotypical findings were explained by the hypothesis that women try to escape *communal* stereotypes by attacking equally (or more frequently) to show voters that they are fit for political roles that are considered *agentic* (Gordon et al., 2003).

Despite these non-stereotypical findings in campaigns, European literature on attacks outside these periods has identified more gender-conforming attack behaviour. More specifically, Hargrave and Langengen

(2021) and Haselmayer et al. (2021) recently looked at differences in speech styles between female and male members of parliament (MPs) in the national parliaments of the UK and Austria, respectively. While controlling for already established predictors, such as the difference between government and opposition, they identify that women employ less adversarial and negative speeches than men. These findings are also in line with Ketelaars (2019), who surveyed Belgian politicians (including members of the parliament) outside campaigns, finding that men prioritise attacking strategies more than women. Therefore, unlike campaigns, these studies corroborate the expectations set by the role congruence theory.

The causes of contrasting behaviour in parliaments and campaigns may be linked to the more versatile approach female politicians are expected to take to achieve their political goals. In other words, female politicians are caught in a double bind between behaving in a *communal* manner (as is expected because they are women) and an *agentic* manner (as is expected because they are politicians). Given that citizens perceive politicians as *agentic*, female politicians need to escape *communal* stereotypes during campaigns by attacking as much as men to secure re-election (Gordon et al., 2003; Maier, 2015). However, in parliaments, politicians compete over policy goals, such as pushing for a specific issue to be high on the agenda (Green-Pedersen & Mortensen, 2010) or trying to acquire ownership over issues (Otjes & Louwerse, 2018). As such, female politicians may evaluate that *communal* behaviour benefits achieving their policy aspirations, while *agentic* behaviour benefits their re-election aspirations.

However, this argument raises the question of why female politicians would conform to *communal* behaviour in parliaments if they already show *agentic* behaviour during campaigns. The cause of this may be due to parliamentary venues traditionally being workplaces that adhere to gender stereotypes (Erikson & Verge, 2022). Therefore, female politicians opting for *communal* behaviour in a dominantly gender-conforming venue such as parliaments provide a greater chance to profile certain policies higher on the agenda or secure their implementation. This is in contrast to campaigning venues, where expectations come from voters who see politics and politicians as *agentic*, which leads to a shift in female politicians' behaviour. Male politicians, in turn, can opt for *agentic* behaviour both in parliaments and campaign venues, as both align with their stereotypical gender roles (parliament) and expectations of them as politicians (campaigns). This is why I hypothesise that women will be less likely to attack in parliaments when compared to men (H1a). However, because of the inevitable elections and the double bind that encourages women to engage in *agentic* behaviour during campaigns, it is expected that the effect of H1a decreases as proximity to the upcoming election increases (H1b).

H1a: Female politicians are less likely to attack compared to male politicians in parliaments.

H1b: The effect of H1a decreases as proximity to elections increases.

Still, if parliaments dominantly represent venues for gender-conforming behaviour to maximise political goals, it is unclear how this translates toward targets of parliamentary attacks. This is why I extend the theoretical framework by arguing that stereotypical gender roles apply not only to the mere decision to attack (or not) but also to a decision of whom to target in an attack. Namely, if most politicians abide by gender-stereotypical behaviour in parliament, with men attacking more than women (H1a), it is also very likely that men are targeted more than women. This decision to attack men more frequently also stems from the role congruency theory, whereby female politicians, due to their association with *communal* roles, are not seen as possible targets of attacks that would otherwise place them in an *agentic* context. Therefore, if an attacker targets a woman, who is not associated with *agentic* norms like men, this may backfire, causing the attacker to be perceived with disapproval because the decision of who to target diverges from expected gender roles (for a somewhat similar claim, see Haselmayer et al., 2021, p. 6). As such, attacking women who are seen as *communal* can be costly for the attacker. This is unlike attacks that target men who are seen as *agentic*, so targeting them and placing them in an agentic framework is expected and can even be beneficial (Fridkin et al., 2009). This is why I argue that female politicians are less likely to be attacked than male politicians in parliaments (H2a). Regardless, given the expectation that behaviour tends to be more *agentic* due to the increasing proximity of the election campaign and vote-seeking goals, it may be that the boomerang effect of attacking female politicians also decreases closer to elections as more attacks are issued. Hence, I argue that the impact of H2a decreases as proximity to elections increases (H2b).

H2a: Female politicians are less likely to be targeted compared to male politicians in parliaments.

H2b: The effect of H2a decreases as proximity to elections increases.

At the same time, there are other predictors for behaviour in parliaments, such as a politician being part of the opposition or the government (Hix & Noury, 2016). We know from the parliamentary literature that the opposition is expected to hold the government accountable (De Giorgi & Ilonszki, 2018). This is because the government holds the keys to the office and has policy perks, which makes it a target of attacks (sometimes even from the majority benches; e.g., Kam, 2009; Martin & Whitaker, 2019). Therefore, depending on their role

in the political system (i.e., cabinet, majority, or opposition), politicians may feel pressured to behave contrary to the gender stereotypes in parliaments. For instance, women in the opposition may be required to be critical and employ *agentic* behaviour. Their role is hence at odds with the *communal* perception expected of them in gender-conforming parliaments, which may hurt their policy goals (H1a). Similarly, female politicians in the government, due to their position, are expected to be targets of attacks. However, because of gender stereotypes, aggressive behaviour towards female cabinet members may backfire (H2a).

This begs the following question: How do politicians balance the costs and the benefits of attacking and being targeted when they diverge from gender stereotypes in parliaments? I expect incivility, seen as a communication interaction that violates social norms (see more in Walter, 2021), to be a possible answer. To appease gender stereotypes, there will be less incivility whenever women do attack or are targeted (H3a/H4a). For example, when the government's policy fails, female politicians in the opposition will likely have to engage in attack behaviour. However, to limit the cost of diverting from the gender stereotype (which may cause prejudice and hurt their goals), female politicians will try to be as polite as possible. In turn, their male colleagues are expected to employ more incivility due to the *agentic* nature of incivility not being costly for them (Bauer et al., 2022; Goovaerts & Turkenburg, 2021). Furthermore, I also expect female targets to be less likely to receive an uncivil attack since campaigning studies show that the presence of women in political debates lowers incivility (Maier & Renner, 2018). This means that all politicians, when forced to target a woman, will restrain from uncivil language. In turn, when targets are males, incivility is more likely to be employed. Lastly, if there is pressure closer to the election to increase non-stereotypical gender behaviour (H1b/H2b), then it is also plausible to expect that the usage of incivility in attacks decreases to compensate for such divergence (H3b/H4b).

H3a: Female politicians are less likely to attack using incivility compared to male politicians in parliaments.

H3b: The effect of H3a increases as proximity to elections increases.

H4a: Female politicians are less likely to be targeted with incivility compared to male politicians in parliaments.

H4b: The effect of H4a increases as proximity to elections increases.

Lastly, while it is expected that there is gender-conforming behaviour in parliamentary venues, there may be differences across different systems (Hargrave & Langengen, 2021, p. 583). This is why I borrow the distinction from the emerging literature on the politics of legislative debate regarding candidate vs party-centred systems (Fernandes et al., 2021). If citizens elect candidates, there is more importance on individual politicians and their own reputations during parliamentary debates (Proksch & Slapin, 2012). However, if citizens elect parties, there is a stronger emphasis on the party brand that diminishes individual characteristics. For example, scholars have shown how in the candidate-driven parliament of the UK, there can be a disconnect between what politicians from the same party feature on their issue agendas (Bevan & John, 2016) with individual politicians focusing on representing their individual constituencies (Blumenau & Damiani, 2021). This is unlike the party-driven parliaments of Belgium, for example, where there is strong party discipline concerning issues that need to be addressed (Peeters et al., 2021).

Because of this, I argue that politicians in candidate-dominated systems are more prone to gender-congruent attack behaviour because there is more emphasis on them as individuals. As such, if female politicians in candidate-driven parliaments divert from stereotypical behaviour, there is much on the line. For example, they may face the consequence of not securing a policy that would be beneficial for their electoral constituency. They may also have to deal with disapproval from the party leadership that may prevent them from seeking re-election in a constituency, especially if there are no gender-related legislative quotas to secure certainty of women re-appearing on ballots. Such a context is unlike party-driven systems where parties provide a certain level of protection from individual gender-incongruent attack behaviour. For example, even if female politicians face the cost of diverting from gender stereotypes in these systems, they can still secure their policy through their party and rely on voters electing their parties, not them individually. This may further be enhanced with gender quotas which would ensure female politicians' spots on a ballot to seek re-election despite diverting from stereotypical gender roles.

H5a: Female politicians adhere more to gender-congruent attack behaviour in candidate-centred compared to party-centred parliaments.

3. Methodology

3.1. Cases

I test my expectations on parliamentary QTSs from the (federal) parliaments of Belgium (*Vragenuur*), Croatia (*Aktualno Prijepodne*), and the UK (*Prime Minister's Questions* [PMQs]). I work with these debates because they present high gain opportunities for politicians to reach their goals due to the heavy media exposure QTSs tend to receive (Osnabrügge et al., 2021; Salmond, 2014). This makes it a perfect case of parliamentary politics to explore whether there are gender differences in attack

strategies that seek to fulfil politicians' goals. This was empirically demonstrated in several studies conducted on QTSs from Belgium (Sevenans & Vliegenthart, 2016; Vliegenthart & Walgrave, 2011), Croatia (Kukec, 2022; Poljak, 2022), and the UK (Bevan & John, 2016; Seeberg, 2020) which have shown how politicians use QTSs to fulfil their policy aspiration, such as placing issues higher on the agenda and trying to pursue voters to elect them at the upcoming elections.

Furthermore, I work with Belgium, Croatia, and the UK because of vast differences in (a) how these QTSs are structured across these three countries and (b) possibilities (and incentives) for female representatives to engage in QTSs. This is important as it allows to test the theory in a robust setting across highly different cases, ensuring a certain level of generalisation while lowering possible selection bias (e.g., studying a specific context of low female representation, which can have implications for parliamentary behaviour; see Sarah & Mona, 2008). Given the importance of these differences, I will reflect on them in greater detail.

Regarding QTS differences, these are highly rooted in the electoral (party) system of each country. Namely, due to the proportional elections where citizens elect parties, the parliaments of Belgium and Croatia are an example of party-driven venues. This party-driven context is reflected in parliamentary procedures where it is parties, and not individual politicians, that are granted slots to ask questions to the cabinet during QTSs (weekly in Belgium; quarterly in Croatia). In Belgium, which can be described as a partitocracy, each major party is granted an equal number of slots during QTSs. In Croatia, which does not have such a strong and stable party system as Belgium, slots during QTSs are granted based on the share of seats. This rule favours two major competing parties in Croatia that employ strong party discipline in QTSs (see Kukec, 2022). As a result, politicians are usually expected to follow party lines during QTSs in both countries. For example, studies from Belgium (De Vet & Devroe, 2022) and Croatia (Šinko & Širinić, 2017) have highlighted how female politicians during (plenary) QTSs tend to profile soft issues, unlike men who deal with hard issues (see also Bäck & Debus, 2019). This is a likely outcome of a strong party discipline during high-profile debates, such as QTSs, where parties select politicians to raise issues that fit their profile (De Vet & Devroe, 2022). While both countries allow preferential voting, this mechanism provides little incentive for politicians to deviate from their parties, as entry to the parliament based on preferential voting is difficult to achieve in both Belgium (Van Erkel & Thijssen, 2016) and Croatia (Picula, 2020). Both countries also have gender quotas that try to ensure that the share of women and men on ballots remains fairly equal, providing a safety net for female parliamentarians already elected to (possibly) re-appear on a party's ballot.

The UK parliament, on the other hand, can best be described as candidate-driven due to the majoritarian elections where citizens elect politicians in single-member districts (Proksch & Slapin, 2012). This doesn't mean that parties are not as important as in Belgium and Croatia, as they still play a major role in getting a politician elected to the parliament (Blumenau & Damiani, 2021, p. 779), and no gender-related legislative quotas are imposed on parties when determining who will run for a party in constituencies. However, once inside the parliament, parties have an incentive to let politicians act in their own personal interest and that of the constituency they represent (Blumenau & Damiani, 2021; Proksch & Slapin, 2012). This is in line with parliamentary procedures as QTSs in the UK (specifically PMQs) are structured by individual questions asked to the prime minister (PM; or a cabinet member when the PM is absent). Only the opposition leaders are granted secure slots to question the PM, while other members who want to question the PM are decided by a random shuffle. This provides less interference from the party leadership and allows politicians to have a certain level of autonomy during QTSs.

Regarding differences in (descriptive) female representation, although all three countries had both male and female PMs, ministers, and party leaders participating in QTSs, the representation of female politicians during QTSs differs vastly (see Table 1). Belgium has a high share of females elected in the parliament, with an average of 39.2% for the last four elections. However, looking at the randomly selected sample of QTSs during the two full parliamentary terms that took place in the 2010s, female politicians were generally underrepresented by nine percentage points in QTSs compared to the share of how many were elected. In turn, Croatia has a significantly lower share of elected female representatives than Belgium (the average for the last five elections is 18%); however, they tend to be overrepresented during QTSs in the last decade. Finally, the UK is somewhere between Belgium and Croatia regarding elected female representatives, with an average of 27% of females elected for the past five elections. Furthermore, unlike in Belgium and Croatia, representation during QTSs in the UK (determined by a random shuffle) generally ensures a fairly equal representation of female MPs during QTSs. As such, with this case selection, we capture parliaments that typically provide lower (Belgium), equal (UK), or higher (Croatia) possibilities for female politicians to participate in QTSs, which makes the chance of selection bias lower than if we had worked with one specific parliamentary setting.

3.2. Speech data during QTSs

To explore attack behaviour and incivility usage longitudinally during QTSs in all three countries, I randomly sampled one QTS per month from January 2010 to December 2020 (2021 for Croatia). This resulted in a total of 261 QTSs in my sample, which covered all quarterly QTSs in Croatia (N = 43; 100%) and 1/3 of all weekly QTSs in

Table 1. The share of women elected to the parliament and the average share of women that participated in QTSs.

Country	Term	Share of elected female politicians *	Average share of female politicians per QTS **	Difference
Belgium	2007–2010	36.7	37.2	+0.5
	2010–2014	39.3	30.4	−8.9
	2014–2018	39.3	30.5	−8.8
	2018–2023	41.3	41.4	+0.1
Croatia	2007–2011	20.9	29.5	+8.6
	2011–2015	19.9	23.4	+3.5
	2015–2016	15.2	16.7	+1.5
	2016–2020	12.6	20.4	+7.8
	2020–2024	23.2	30.4	+7.2
UK	2005–2010	19.8	18.1	−1.7
	2010–2015	22.0	23.3	+1.3
	2015–2017	29.4	28.9	−0.5
	2017–2019	32.0	34.8	+2.8
	2019–2024	33.8	28.5	−5.3

Notes: * Based on: Institute for the Equality of Women and Men (2022) for Belgium (*Chamber of Representatives*); Šinko (2016) for Croatia; Allen (2020) for the UK (*House of Commons*). ** Indicates average share of females that participate per QTS in the sampled period (N = 261; see Section 3.2). More detailed descriptive results are available in Appendix C in Supplementary File 1.

Belgium (N = 103; 30.4% out of all QTSs) and the UK (N = 115; 32.7%). To generate raw data on individual politicians who attacked and were targeted within these QTSs (and whether incivility was involved), I scraped transcripts of sampled QTSs from official parliamentary websites where units of observation constitute every speech contribution during the sessions. Protocol speeches such as speakers moderating the debate (only in Croatia) or the UK's PMs listing their engagements at the start of every PMQ were not included in this data (Belgium N = 6,634; Croatia N = 5,679; UK N = 7,731).

Four coders, together with the author, performed a quantitative content analysis on these speech contributions in which the main goal was to collect information on attacks (a six-week training process with reliability scores is presented in Appendix A in Supplementary File 1). Coders had to carefully read each speech contribution during QTSs and identify (a) if an attack was present or absent. The codebook defines attacks following Geer (2006) as *all criticism directed toward political competitors* but also extends this to internal attacks as parties are prone to intra-party dissent in parliaments (Kam, 2009), and coalition partners may hold each-other accountable (Martin & Whitaker, 2019). Therefore, attacks can only be directed towards formal political actors seen as individuals (e.g., PM, Ministers), parties (e.g., Conservatives, Labour), or groups of parties (e.g., opposition, coalition, government). Attacks directed towards informal actors (e.g., the army, NGOs, foreign actors) are not coded.

If a speech unit was coded as containing an attack, coders proceeded to code (b) if there was a presence of incivility. As stated in the theory, incivility is operationalised as a communicative interaction between political actors that violates social norms (Walter, 2021) and is therefore nested in attacks (i.e., attacks can either have incivility present or absent). As such, any form of name-calling, mocking, or insulting language used in attacks on politicians and their policies was coded as incivility. Lastly, coders also had to indicate who was on the receiving end of the attack, and in the case of multiple attacks, coders coded each attack separately. As such, in one speech unit, one actor may have attacked both Minister A and B, so coders had to indicate for both targets separately whether they were attacked with incivility or not. In total, 6,643 speeches or 33.2% had at least one attack present (Belgium 32.7%; Croatia 36.8; UK 30.9%) and the overall number of attacks was 9,485 (Belgium N = 3,117; Croatia N = 3,339; UK N = 3,029). 1,735 or 18.3% of all attacks made were employed using incivility (Belgium 23.5%; Croatia 15.4%; UK 16.1%). Examples of civil, uncivil, and non-attacks per country are available in Appendix B in Supplementary File 1.

3.3. Final Data

Based on coded speeches, I was able to generate new data to test hypotheses. This data includes all individual politicians that participated during QTSs (Belgium N = 369; Croatia N = 468; UK N = 833) which are observed per each QTS in which they made at least one speech contribution (Belgium N = 2,898; Croatia N = 2,354; UK N = 2,930). As such, on a QTS where 40 politicians spoke (e.g., by asking questions, giving answers, raising points of order), each of these 40 was treated as a unique observation for that particular QTS (see Table 2). This

allows me to explore whether these politicians decided to employ an attack (dependent variable 1 [DV1]) and/or were targeted (dependent variable 2 [DV2]) during that particular QTS. In turn, when values in these two variables are 1, it indicates that an actor employed an attack and/or was targeted; data also indicates if incivility was present in any attacks that were employed (dependent variable 3 [DV3]) or received (dependent variable 4 [DV4]). These four constitute binary dependent variables of my study, each of which corresponds to the four hypotheses, while gender (male vs. female) and proximity to elections (i.e., how many months have passed since the last parliamentary election) present the main independent variables. Speakers that moderate QTSs are omitted because they are bound to attack regularly on QTSs when rules of procedures are not followed.

Four control variables are included in the data: the politician's position (opposition, majority, or cabinet), country, ideology, and inter-annual (yearly) dummies. Ideology is generated using Chapel Hill Expert Survey (CHES) data (Jolly et al., 2022), where the average ideological scores of each party in the studied period are attributed to their respective members. These scores are then calculated for divergence from the political centre, with 0 indicating the political centre. As such, the bigger the score, the more ideologically extreme politicians are. Descriptive statistics for variables are available in Appendix C in Supplementary File 1.

3.4. Method

I employ logistic regressions due to the binary nature of my DVs. These regressions are run through multi-level models because data is hieratical, with politicians being observed on two levels: parties (N = 39) and QTSs (N = 261). Both levels are entered as random intercepts in the model in which the level of parties is crossed in the level of QTSs in which they appear (Figure 1). This (multiple-membership) multi-level modelling strategy is important because it accounts for the fact that politicians

Table 2. Example of the final dataset.

N	Date	Election proximity	Politician	Party	Gender	Employing attack	Being targeted	Attacking with incivility	Being targeted with incivility
7,954	13.2.2019	20	Theresa May	Con	1	1	1	1	1
7,955			Helen Whately	Con	1	0	0	—	—
7,956			Craig Mackinlay	Con	0	1	0	0	—
7,957			Jeremy Corbyn	Lab	0	1	1	1	0
7,958			Vicky Foxcroft	Lab	1	0	0	—	—
7,959			Luke Pollard	Lab	0	1	0	0	—
7,960			Liz Saville Roberts	PC	1	0	1	—	0
7,961			Ian Blackford	SNP	0	1	1	1	0
7,962			Mhairi Black	SNP	1	0	0	—	—
(...)			(...)	(...)	(...)	(...)	(...)	(...)	(...)

Note: Only a fraction of data is shown for one QTS in the UK.

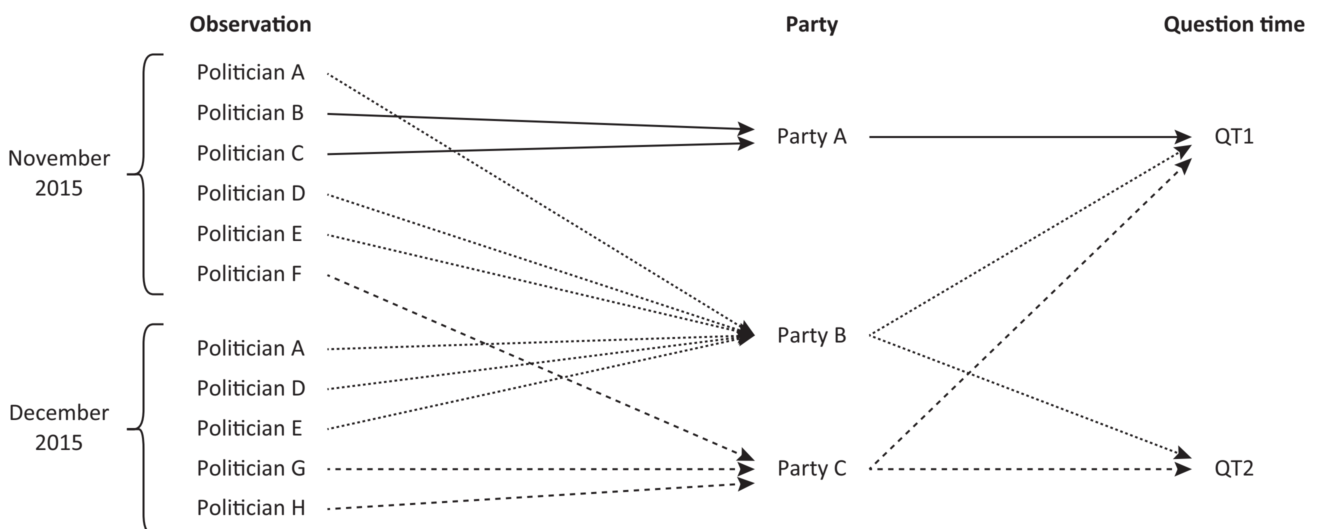

Figure 1. Multi-level model observing politicians per each party crossed in QTSs.

of each party re-appear as observations throughout my data. As such, this modelling strategy prevents biases where politicians from a certain party may skew the results of the model, while in reality, they all belong to one party that re-appears across the dataset (see Chung & Beretvas, 2012). When running these models, I drop all politicians who are independent or whose parties are not featured in the CHES dataset.

4. Results

I will first test my five main hypotheses (H1a–5a), after which I will explore trends as the proximity to parliamentary elections increases (H1b–4b). The results from my main models are reported in Table 1 and Figure 2. As can be seen, the results show support for H1a, H2a, H3a, and H4a (for descriptive analyses, see Appendix D in Supplementary File 1). Odds that female politicians will attack (H1a), be targeted (H2a), use incivility when they attack (H3a), and be targeted with incivility (H4a) during QTSs significantly decrease when compared to their male colleagues. Overall, these multivariate analyses show strong support for the main theory of this article on how women and men behave according to their stereotypical gender roles in parliamentary attacks. Furthermore, when women need to attack, such as when they are in the opposition, we can expect them to be less likely to employ incivility. Lastly, when women are at the receiving end of attacks, there is a greater chance that these attacks will be civil, unlike those that target men.

To test H5a, that there are also differences among countries, I run models that interact variables on gender and country. For H1a, H2a and H3a, there is a significant difference across countries, with women conforming to gender expectations more in the UK when compared to Belgium and Croatia (see regressions' output in Appendix F in Supplementary File 1). In addition to that, running models separately in each country further confirms this. While coefficients in almost every model go in a negative direction (with lower odds of women engaging in attacks and incivility than men), these are significant in the UK but less so in Belgium and Croatia. Specifically, in Belgium, I can reject all

Table 3. Multi-level regressions testing probabilities of engaging in attacks during QTSs.

	Model 1	Model 2	Model 3	Model 4
	DV1: Employing attack (1 = Yes)	DV2: Being targeted (1 = Yes)	DV3: Attacking with incivility (1 = Yes)	DV4: Being targeted with incivility (1 = Yes)
	Coef. (S.E.)	Coef. (S.E.)	Coef. (S.E.)	Coef. (S.E.)
Male politicians (ref.)				
Female politicians	**−.210 (.062) ****	**−.405 (.079) *****	**−.473 (.101) *****	**−.312 (.144) ***
Proximity to Elections	.013 (.002) ***	.007 (.002) **	.008 (.003) **	.008 (.005) †
Ideology	1.493 (.855) †	.358 (1.072)	3.072 (.887) **	.056 (1.072)
Opposition MPs (ref.)				
Majority MPs	−2.869 (.102) ***	−3.082 (.158) ***	−1.129 (.157) ***	−.418 (.326)
Cabinet politicians	−2.497 (.107) ***	1.112 (.103) ***	.140 (.148)	.914 (.155) ***
Belgium (ref.)				
Croatia	.382 (.221) †	.568 (.279) *	−.737 (.259) **	−.331 (.247)
UK	−1.028 (.253) ***	.203 (.338)	.126 (.277)	.365 (.279)
Constant	.908 (.291) **	−1.532 (.345) ***	−1.390 (.327) ***	−1.495 (.401) ***
Variance (QTSs)	.364 (.041)	.083 (.146)	.222 (.093)	.260 (.136)
Variance (Parties)	.429 (.096)	.617 (.106)	.430 (.097)	.307 (.129)
N (total)	7,724	7,724	3,140	1,595
N (QTSs)	261	261	261	261
N (min. politicians per QTS)	13	13	3	1
N (max. politicians per QTS)	56	56	37	23
AIC (empty model)	8.140 (9.509)	5.785 (7.707)	3.584 (3.810)	1.938 (1.984)

Notes: †p < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001; Control for yearly differences included.

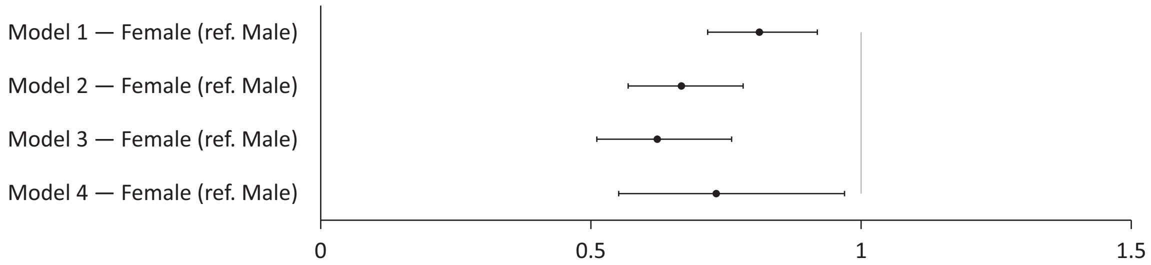

Figure 2. The odds ratio of the gender variable calculated from the models in Table 3. Notes: Ratios < 1 indicate lower odds of women appearing in an attack compared to men; Horizontal lines indicate confidence intervals (95%); Odds ratio for all variables available in Appendix E in Supplementary File 1.

hypotheses. In Croatia, I find support for H2a while H3a is relatively close to being significant (p = 0.071). In contrast to these two countries, H1a, H2a, H3a, and H4a have support in the UK. As such, there is a strong indication that H5a holds and that gender-conforming behaviour is more visible in the candidate-driven compared to the party-driven parliaments.

Finally, I test H1b-H4b, which argued that women's adherence to gender stereotypes decreases as proximity to the upcoming election increases while the protection mechanism of not using incivility increases. Given the null findings of gender-conforming attack behaviour in Belgium and Croatia, I specifically focus on the case of the UK to test these expectations. Namely, I run models that interact variables on gender and proximity to elections, after which I plot predicted probabilities of these interactions to inspect patterns of attack behaviour throughout the electoral cycle (regressions' output and tests for Belgium and Croatia are available in Appendix G in Supplementary File 1).

As is demonstrated in Figure 3, there is mixed support for H1b and H2b. Namely, gender-conforming behaviour expected from H1a and H2a exists regardless of elections, with women attacking and being targeted significantly less than men throughout the UK electoral cycle. Still, comparing increases in average probabilities throughout the electoral cycle for men and women separately, we can descriptively confirm certain differences. For example, when comparing the first month after an election to the final month before an election, the average probability of an attack being employed increases by 33% for women (from 0.2 to 0.27) and 15.3% for men (from 0.3 to 0.34). As such, on a descriptive level, women do increase their attack behaviour closer to elections more strongly when compared to men. This is likely a result of the double-bind argument in which women have to balance both *communal* and *agentic* norms through time. This is unlike men who can opt for agentic behaviour regardless of elections, making their increase in attack behaviour less profound.

Moving to incivility usage in attacks, we see that women use incivility less often than men, regardless of the electoral cycle in the UK. However, as is visible in Figure 4, there is merit to H3b. Specifically, closer to elections, as women diverge from gender stereotypical roles by increasing attack behaviour (Figure 3), they also try to increase the protection of such divergence by lowering their usage of incivility. This is unlike men whose incivility increases closer to elections. For example, when the first month after an election is compared to the final month before an election, the average probability of incivility being used in an attack decreases by

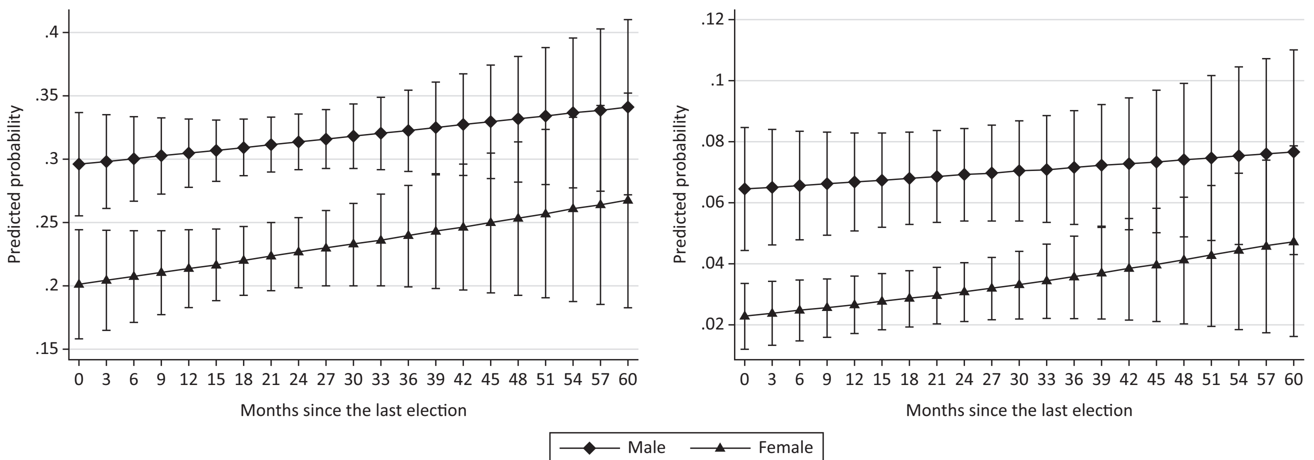

Figure 3. Predicted probabilities for employing attack (left) and being targeted (right) during QTSs in the UK. Note: Vertical lines indicate confidence intervals (90%).

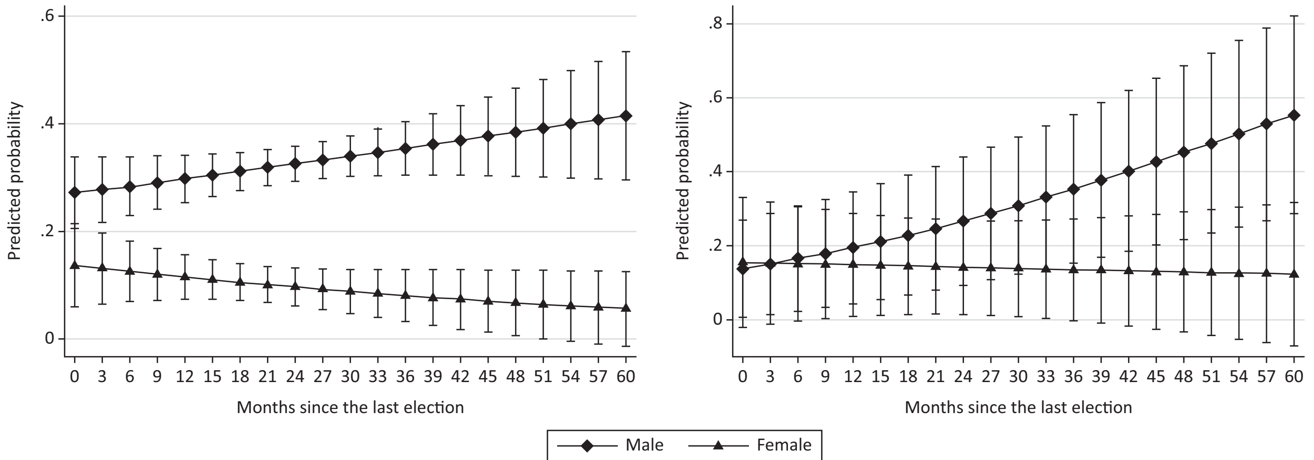

Figure 4. Predicted probabilities for employing attack with incivility (left) and being targeted with incivility (right) during QTSs in the UK. Note: Vertical lines indicate confidence intervals (90%).

57.1% for women (from 0.14 to 0.06) while it increases by 51.9% for men (from 0.27 to 0.41). In turn, regarding H4b, results show how women can be targeted with incivility equally to men, but the increase in receiving uncivil attacks closer to elections is more profound for men, which is in line with H4b. Overall, while expectations regarding H1b–H4b are not confirmed on a level of statistical significance (Appendix G in Supplementary File 1), the evolution of attack behaviour throughout the electoral cycle demonstrated in Figures 3 and 4 shows that there is some ground for the hypothesised outcomes in the UK. This is especially true regarding H3b, with women decreasing and men increasing incivility as the overall attack behaviour increases closer to elections.

4.1. Robustness Checks

To ensure the validity of the results, all binary DVs have been transformed to count DVs that indicate the total number of attacks. Negative binomial regressions are run, and the results corroborate findings from the multilevel logistic regressions (Appendix H in Supplementary File 1). In the UK, the theory shows strong support for men employing more and receiving more attacks than women. In Belgium, H1a is close to statistical significance ($p = 0.053$), revealing that male politicians in Belgium likely employ more attacks than women. However, for other hypotheses, no support exists, and the same applies to Croatia, where all hypotheses can be rejected using count DVs.

However, to further strengthen the findings that gender differences drive the attack behaviour of politicians in the UK parliament, I run further tests (Appendix I in Supplementary File 1). Namely, I explore the seniority of MPs (i.e., years since the first entry to the parliament) and also their position in the parliament (frontbenchers such as PMs, Cabinet Members, Opposition Leaders, Shadow Ministers, Party Leaders, and Parliamentary Group Leaders, vs. backbenchers who do not hold any official role in a party or the parliament). Adding these controls to the main models further confirms that it is indeed female politicians who are significantly less likely to attack (H1a), and that when they do, they will be significantly less likely to use incivility (H3a). However, adding control for the position (frontbench vs. backbench) does diminish findings regarding targets (H2a/H4a); compared to backbenchers, frontbenchers receive more attacks, which are more likely to be uncivil in nature.

5. Conclusions

This study contributes to the current negativity literature by providing an overarching theoretical framework that provides us with an understanding of parliamentary attacks from a gender perspective. Namely, in candidate-driven parliaments such as the one in the UK, we can expect attacks to be conditioned on gender, with female politicians attacking less frequently. However, given that female politicians are caught in a double bind by trying to appease expectations of being a woman and a politician, their behaviour during the term is likely to change. As the time during the cycle elapses, women increase agentic behaviour by employing more attacks which may grant them re-election. In turn, while employing more attacks, women lower their usage of incivility as they are likely trying to mitigate possible costs for their divergence from stereotypical gender expectations. This behaviour is distinct from male politicians, who also increase attacks during the term, but their incivility usage increases closer to elections as they face less cost for displaying agentic behaviour. On the other hand, in party-driven parliaments such as those in Belgium and Croatia, we can expect politicians not to conform to stereotypical gender behaviour. Safe in the knowledge that they can rely on their parties to feature issues high on the agenda or acquire ownership of certain issues (which in the long run provides more possibility for re-election through parties), female parliamentarians have greater freedom to

not adhere to gender stereotypes regarding attacks and the use of incivility.

Besides contributing to the negativity literature, this study also contributes to the gender literature on female representation. Despite differences in female (descriptive) representation in the parliaments of Belgium and Croatia, in both cases, female representatives behave similarly by not conforming to gender expectations regarding attacks. In contrast, gender-conforming attack behaviour is present in the UK. As such, we can align with the scholarly work that has also found limited support for different patterns of female parliamentary behaviour if the proportion of women in parliaments changes (Sarah & Mona, 2008, p. 733). This study highlights the importance of the broader institutional setting (see Lovenduski, 2019) when it comes to studying the political behaviour of politicians based on gender. Therefore, different attack behaviour between men and women across the countries may be rooted in the electoral systems and the different possibilities of securing policy goals and re-election; in Belgium and Croatia, politicians act within and in the interest of their parties supported by gender quotas, whereas in the UK politicians act individually and in the interest of their constituencies, without the security provided by gender quotas.

However, while it is likely that the peculiar exception of the UK is an outcome of its candidate-driven parliamentary system, whereby individuals are more prominent in issue and party competition, it is important to reflect on the limitation that this finding comes from one particular case. In other words, it may be that these peculiar findings of gender-conforming attack behaviour are more likely in the context of UK politics and not necessarily in systems where individuals also seek re-election in single-member districts. As such, given this study's limitations, it is important to investigate whether the findings from the UK apply to other parliamentary systems that are candidate-oriented to ensure the generalisability of the theory. Yet, given the similarities regarding the treatment of female politicians across Westminster-style parliaments (e.g., Collier & Raney, 2018), there are reasons to suspect that findings may be applicable beyond the UK case. Furthermore, this study only focused on a specific format of parliamentary politics (QTSs), neglecting all other forms of debates such as committee sittings. Therefore, future studies should dive deeper into the mechanisms that possibly shape attack behaviour in other parliamentary debates. Lastly, future studies should also explore the content of attacks, which may uncover currently neglected patterns of attack. It may be, for example, that women attack equally to men in Croatia and Belgium, but the content of their criticism might differ vastly.

Acknowledgments

I am grateful to anonymous reviewers who provided valuable feedback on the previous version of this article. The article also benefitted greatly from comments provided by Julie Sevenans, Stefaan Walgrave, Annemarie Walter, and Peter Van Aelst. I acknowledge financial support from the University of Antwerp Research Fund (BOF) (grant number 43834).

Conflict of Interests

The author declares no conflict of interest.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Allen, G. (2020). *General Election 2019: How many women were elected?* House of Commons Library. https://commonslibrary.parliament.uk/general-election-2019-how-many-women-were-elected

Auter, Z. J., & Fine, J. A. (2016). Negative campaigning in the social media age: Attack advertising on Facebook. *Political Behavior*, *38*(4), 999–1020.

Bäck, H., & Debus, M. (2019). When do women speak? A comparative analysis of the role of gender in legislative debates. *Political Studies*, *67*(3), 576–596.

Banwart, M. C., & Bystrom, D. G. (2022). Gender and candidate communication: An analysis of televised ads in the 2020 US Senate Races. *American Behavioral Scientist*. Advance online publication. https://doi.org/10.1177/00027642211040766

Bauer, N. M., Kalmoe, N. P., & Russell, E. B. (2022). Candidate aggression and gendered voter evaluations. *Political Psychology*, *43*(1), 23–43.

Bevan, S., & John, P. (2016). Policy representation by party leaders and followers: What drives UK Prime Minister's Questions? *Government and Opposition*, *51*(1), 59–83.

Blumenau, J., & Damiani, R. (2021). The United Kingdom: The (increasing) discretion of MPs in parliamentary debate. In H. Bäck, M. Debus, & J. M. Fernandes (Eds.), *The politics of legislative debates* (pp. 775–800). Oxford University Press.

Bystrom, D. G. (2004). Women as political communication source and audiences. In L. L. Kaid (Ed.), *Handbook of political communication research* (pp. 435–459). Routledge.

Chung, H., & Beretvas, S. N. (2012). The impact of ignoring multiple membership data structures in multilevel models. *British Journal of Mathematical and Statistical Psychology*, *65*(2), 185–200.

Collier, C. N., & Raney, T. (2018). Understanding sexism and sexual harassment in politics. *Social Politics: International Studies in Gender, State & Society*, *25*(3), 432–455.

De Giorgi, E., & Ilonszki, G. (Eds.). (2018). *Opposition parties in European legislatures*. Routledge.

De Vet, B., & Devroe, R. (2022). Party control, intra-party competition, and the substantive focus of women's parliamentary questions: Evidence from Belgium. *Politics & Gender*. Advance online publication. https://doi.org/10.1017/S1743923X21000490

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological review*, *109*(3), 573–598.

Erikson, J., & Verge, T. (2022). Gender, power and privilege in the parliamentary workplace. *Parliamentary Affairs*, *75*(1), 1–19.

Evans, H. K., Cordova, V., & Sipole, S. (2014). An analysis of how house candidates used Twitter in their 2012 campaigns. *PS: Political Science & Politics*, *47*(2), 454–462.

Fernandes, J. M., Debus, M., & Bäck, H. (2021). Unpacking the politics of legislative debates. *European Journal of Political Research*, *60*(4), 1032–1045.

Fridkin, K. L., Kenney, P. J., & Woodall, G. S. (2009). Bad for men, better for women: The impact of stereotypes during negative campaigns. *Political Behavior*, *31*(1), 53–77.

Geer, J. G. (2006). *In defense of negativity: Attack ads in presidential campaigns*. The University of Chicago Press.

Goovaerts, I., & Turkenburg, E. (2021, May 27–31). *It's the context, stupid! Investigating patterns and determinants of political incivility in televised debates over time* [Paper presentation]. 71st Annual ICA Conference, Virtual Conference.

Gordon, A., Shafie, D. M., & Crigler, A. N. (2003). Is negative advertising effective for female candidates? An experiment in voters' uses of gender stereotypes. *Harvard International Journal of Press/Politics*, *8*(3), 35–53.

Green-Pedersen, C., & Mortensen, P. (2010). Who sets the agenda and who responds to it in the Danish parliament? *European Journal of Political Research*, *49*(2), 257–281.

Hargrave, L., & Langengen, T. (2021). Do men and women communicate differently in the house of commons? *Politics & Gender*, *17*(4), 580–606.

Haselmayer, M., Dingler, S. C., & Jenny, M. (2021). How women shape negativity in parliamentary speeches—A sentiment analysis of debates in the Austrian parliament. *Parliamentary Affairs*. Advance online publication. https://doi.org/10.1093/pa/gsab045

Herrnson, P. S., & Lucas, J. C. (2006). The fairer sex? Gender and negative campaigning in US elections. *American Politics Research*, *34*(1), 69–94.

Hix, S., & Noury, A. (2016). Government-opposition or left-right? The institutional determinants of voting in legislatures. *Political Science Research and Methods*, *4*(2), 249–273.

Institute for the Equality of Women and Men. (2022). *Aanwezigheid van vrouwen in de belgische wetgevende en uitvoerende instellingen* [Presence of women in the Belgian legislative and executive institutions]. https://igvm-iefh.belgium.be/nl/activiteiten/politiek/cijfers

Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., Steenbergen, M., & Vachudova, M. A. (2022). Chapel Hill Expert Survey trend file, 1999–2019. *Electoral Studies*, *75*, Article 102420.

Kam, C. J. (2009). *Party discipline and parliamentary politics*. Cambridge University Press.

Ketelaars, P. (2019). Position, preference and personality: A microlevel explanation of negativity in day-to-day politics. *Political Psychology*, *40*(5), 1019–1038.

Kukec, M. (2022). Ask me something I know: Cabinet members in Question Time. *The Journal of Legislative Studies*. Advance online publication. https://doi.org/10.1080/13572334.2022.2085459

Lovenduski, J. (2019). Feminist reflections on representative democracy. *The Political Quarterly*, *90*, 18–35.

Maier, J. (2015). Do female candidates feel compelled to meet sex-role expectations or are they as tough as men? In A. Nai & A. Walter (Eds.), *New perspectives on negative campaigning* (pp. 129–146). ECPR Press.

Maier, J., & Nai, A. (2021). Mapping the drivers of negative campaigning. *International Political Science Review*. Advance online publication. https://doi.org/10.1177/0192512121994512

Maier, J., & Renner, A. M. (2018). When a man meets a woman. *Political Communication*, *35*(3), 433–449.

Martin, S., & Whitaker, R. (2019). Beyond committees: Parliamentary oversight of coalition government in Britain. *West European Politics*, *42*(7), 1464–1486.

Nai, A., & Maier, J. (2018). Perceived personality and campaign style of Hillary Clinton and Donald Trump. *Personality and Individual Differences*, *121*, 80–83.

Osnabrügge, M., Hobolt, S. B., & Rodon, T. (2021). Playing to the gallery: Emotive rhetoric in parliaments. *American Political Science Review*, *115*(3), 885–899.

Otjes, S., & Louwerse, T. (2018). Parliamentary questions as strategic party tools. *West European Politics*, *41*(2), 496–516.

Peeters, J., Van Aelst, P., & Praet, S. (2021). Party ownership or individual specialization. *Party Politics*, *27*(4), 692–703.

Picula, B. (2020, July 20). *Ravno u Sabor s mizerijom od glasova? Može, pita se stranka, a ne birači* [Straight to the parliament with a misery of votes? It is possible, it's up to the party, not the voters]. *Tportal.hr*. https://www.tportal.hr/komentatori/clanak/ovo-je-apsurdno-ravno-u-sabor-s-mizerijom-od-glasova-moze-pita-se-stranka-a-ne-biraci-foto-20200720

Poljak, Ž. (2022). Negative interplay between cabinet and opposition agendas: Evidence from Croatia. *Parliamentary Affairs*. Advance online publication. https://doi.org/10.1093/pa/gsab067

Proksch, S. O., & Slapin, J. B. (2012). Institutional foundations of legislative speech. *American Journal of Political Science*, *56*(3), 520–537.

Salmond, R. (2014). Parliamentary question times: How

legislative accountability mechanisms affect mass political engagement. *The Journal of Legislative Studies*, *20*(3), 321–341.

Sarah, C., & Mona, L. K. (2008). Critical mass theory and women's political representation. *Political studies*, *56*(3), 725–736.

Seeberg, H. B. (2020). The impact of opposition criticism on the public's evaluation of government competence. *Party Politics*, *26*(4), 484–495.

Sevenans, J., & Vliegenthart, R. (2016). Political agenda-setting in Belgium and the Netherlands: The moderating role of conflict framing. *Journalism & Mass Communication Quarterly*, *93*(1), 187–203.

Šinko, M. (2016). Parlamentarna predstavljenost žena u Hrvatskoj [Parliamentary representation of women in Croatia]. *Političke analize*, *7*(27), 3–10.

Šinko, M., & Širinić, D. (2017, October 24–26). *In search of a women-friendly legislative agenda: Case study of Parliamentary Questions in Croatia* [Paper presentation]. International Conference on Gender and Politics 2017, Ouarzazate, Morocco.

Van Erkel, P. F., & Thijssen, P. (2016). The first one wins: Distilling the primacy effect. *Electoral Studies*, *44*, 245–254.

Vliegenthart, R., & Walgrave, S. (2011). Content matters: The dynamics of parliamentary questioning in Belgium and Denmark. *Comparative Political Studies*, *44*(8), 1031–1059.

Wagner, K. M., Gainous, J., & Holman, M. R. (2017). I am woman, hear me tweet! Gender differences in Twitter use among congressional candidates. *Journal of Women, Politics & Policy*, *38*(4), 430–455.

Walter, A. S. (2013). Women on the battleground. *Journal of Elections, Public Opinion & Parties*, *23*(2), 154–176.

Walter, A. S. (2021). The study of political incivility and its challenges. In A. Walter (Ed.), *Political incivility in the parliamentary, electoral and media arena* (pp. 1–16). Routledge.

About the Author

Željko Poljak is a PhD candidate and a member of the M2P (Media, Movements & Politics) research group at the Department of Political Science, University of Antwerp. His current research deals with negative communication between political actors both outside and during election campaigns. His recent work on this topic has been published in *European Journal of Political Research* and *West European Politics*.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 299–310
https://doi.org/10.17645/pag.v10i4.5719

Article

The Personality Origins of Positive and Negative Partisanship

Alexa Bankert

Department of Political Science, University of Georgia, USA; alexa.bankert@uga.edu

Submitted: 29 April 2022 | Accepted: 16 August 2022 | Published: 30 December 2022

Abstract
Negative partisanship describes the intense disdain for a rival political party. A growing number of political scientists in the US and beyond examine the impact of negative partisanship on citizens' political behavior, asserting the notion that negative partisanship exerts a strong influence, either on its own or in combination with positive partisanship. Yet we know little about the psychological origins of negative and positive partisanship: Which personality traits are associated with high levels of negative partisanship, and do they differ from the ones that have been linked to positive partisanship? In this article, I address these questions. Utilizing a sample of US adults and a sample of Swedish adults, I examine the influence of prominent personality traits—including Authoritarianism, Social Dominance Orientation, the Need for Closure, and the Big Five—on strong negative and positive partisanship. I demonstrate that the personality origins of positive and negative partisanship differ not just across the two samples but also across partisans on the left and on the right. I conclude the article with implications for research on polarization and a plea for more comparative work on (positive and negative) partisanship.

Keywords
negative partisanship; personality; positive partisanship; psychology; social identity; Sweden; US

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Negative partisanship (NPID) has received increasingly more attention from scholars in the past few years. While a simple Google Scholar search yields only 1,600 hits for the term "negative partisanship," this number cannot convey the exponential growth of scholarship over the past few years. Indeed, 83% of current research on NPID has been contributed since 2018. Thus, NPID is gaining traction among political scientists who study its effects on political behavior, either in combination with positive partisanship (PPID; Abramowitz & Webster, 2016; Bankert, 2021) or, increasingly, on its own (Caruana et al., 2015; Garzia & Ferreira da Silva, 2022; Mayer, 2017; Medeiros & Noël, 2014; Meléndez & Kaltwasser, 2021; Rose & Mishler, 1998), demonstrating that both PPID and NPID shape vote choice, turnout, and various other forms of political participation (Samuels & Zucco, 2018; Tsatsanis et al., 2020). Despite their demonstrated impact on political behavior, we know little about the origins of PPID and NPID. While prior work examines the role of party leaders (e.g., elite-level polarization) as well as institutional features (e.g., the two-party system), I focus on the individual, psychological origins of NPID and PPID that address several important questions: Who is more likely to develop strong NPID? What kind of personality traits are associated with high levels of NPID, and do they differ from the ones that have been linked to PPID?

To tackle these questions, I examine a range of prominent personality traits, including the Need for Closure (NfC), Authoritarianism, Social Dominance Orientation (SDO), and the Big Five. I focus on these traits because of prior work that has provided evidence for their association with stronger partisan attachments. Yet there is little systematic evidence that examines more than just one

personality trait at a time or that compares their impact on both PPID and NPID. Moreover, there is little prior comparative work on the relationship between personality and partisanship despite the concern that personality traits and their impact might vary across countries due to cultural as well as political differences. To address this gap, I examine a sample of American and Swedish citizens in this study. The US two-party system differs dramatically from the Swedish coalitional multi-party system. Indeed, Sweden is not just characterized by proportional representation but also by a "fundamentally consensual political culture" and much lower levels of affective polarization across partisans (Oscarsson et al., 2021, p. 5). These features stand in sharp contrast to the American two-party system, thereby allowing for a comparison of the personality origins of positive and negative partisan identity in two vastly different political systems.

Last, in contrast to prior work, I utilize a measure that conceptualizes NPID as a social identity rather than just a negative affect towards the out-party (see Abramowitz & Webster, 2018) or a negative vote (see Caruana et al., 2015). While I do not claim this measure to be better in any way than prior measurement strategies, it does feature a few promising measurement properties, such as a multi-item index that can gauge even subtle variations in negative partisan identity strength, high reliability, as well as good predictive power (see Bankert, 2021)—similar to the positive partisan identity scale (see Bankert et al., 2017).

Using these identity scales, I demonstrate that PPID and NPID have very different psychological origins among partisans in both Sweden as well as the US. I also find significant differences across the ideological aisle whereby NPID and PPID on the left are associated with different traits than NPID and PPID on the right. From this perspective, this article contributes a few insights to the contemporary literature on PPID and NPID: First, it provides a systematic and comprehensive overview of the effects of personality traits on strong PPID and NPID. Second, it compares these effects across two vastly different political systems, thereby providing insights into their generalizability. Third, it compares these effects across the ideological left and the right, which elucidates the different psychological compositions of partisan groups.

In the remainder of this article, I first provide a brief overview of the existing literature on personality and partisanship. I then introduce the data and the measurement strategies, including the decision to measure PPID and NPID as an identity. In the analysis part of the article, I examine four different types of partisans: Negative Partisans, who display high levels of NPID but low levels of PPID; Positive Partisans, who display high levels of PPID but low levels of NPID; Closed Partisans, who display both high levels of NPID and PPID; and last, Apathetic Partisans, who display both low levels of NPID and PPID. This comparison clearly identifies the different origins of PPID and NPID as well as the traits that contribute to their alignment. I conclude the article with implications for research on polarization and a plea for more comparative work on (positive and negative) partisanship.

2. Partisanship as a Social Identity

In this study, I consider partisanship a "social identity," which is defined as "that part of an individual's self-concept which derives from his knowledge of his membership in a social group (or groups) together with the value and emotional significance attached to that membership" (Tajfel, 1978). I measure this identity-based conceptualization of partisanship (Green et al., 2004; Greene, 1999, 2002, 2004; Huddy et al., 2015) with a multi-item index which is broadly based on Mael and Tetrick's (1992) "identification with a psychological group scale." With items such as "When I meet somebody who supports this party, I feel connected" and "When people praise this party, it makes me feel good," the scale captures crucial social identity ingredients such as partisans' subjective sense of belonging to the group as well as the importance of the group membership. Since the scale gauges affirmative identification with a political party, I refer to it as the positive partisan identity (PPID) scale.

This social identity framework is also useful for deriving an understanding of NPID. According to social identity theory (Tajfel & Turner, 1979), identities can also form in opposition to groups to which we do not belong. Thus, the identity is negative in the sense that it centers on the rejection of an out-group and its members (Zhong, Galinsky, & Unzueta, 2008; Zhong, Phillips, et al., 2008). In the political arena, Americans form negative identities in response to third parties (Bosson et al., 2006) as well as political organizations like the National Rifle Association (Elsbach & Bhattacharya, 2001), turning the exclusion from a group—the "not being one of them"—into a meaningful social identity. I argue that the same can be true for political parties, whereby the strong rejection of a political party can develop into a negative partisan identity (NPID). In my prior work, I designed and validated a multi-item scale that measures this NPID (see Bankert, 2021). To make the PPID and NPID scales as comparable as possible, I flip the items of the PPID scale to capture the emotional significance respondents associate with their rejection of the out-party with items such as "When I meet somebody who supports this party, I feel disconnected" and "I get angry when people praise this party." The PPID and NPID scale items are listed in Table A1 in the Supplementary File.

3. The (Un-) Alignment of Positive and Negative Partisanship

With two separate measures to capture PPID and NPID, it is possible to create a typology of partisans that can be distinguished by their different PPID and NPID

levels. Early work by Rose and Mishler (1998) has already done so with the example of post-communist countries, whereby the authors examine four different types of partisans: (a) Open Partisans with PPID toward their in-party and without NPID toward another party, (b) Negative Partisans with NPID and without PPID, (c) Closed Partisans with both NPID and PPID, and (d) Apathetic Partisans with no identification. Rose and Mishler find that in the four countries they studied, namely Hungary, Poland, Romania, and Slovenia, more than half of respondents held NPID towards at least one party but PPID towards none. Similarly, Mudde and Kaltwasser (2018) note that in many Western European democracies, populist parties are the targets of NPID despite declining levels of PPID. These findings show that NPID and PPID do not always occur together.

While I partly rely on Rose and Mishler's terminology in this study, I slightly alter their typology. Rather than examining whether a partisan has a positive and/or negative identification with a political party, I examine the intensity or strength of that positive and/or negative identification. This leads to four different types of partisans: Positive Partisans with high levels of PPID and low levels of NPID, Negative Partisans with high levels of NPID and low levels of PPID, Closed Partisans with high levels of both PPID and NPID, and Apathetic Partisans with low levels of both PPID and NPID (see Figure 1). In the next section, I will utilize this typology to make predictions about the distinct personality traits that are associated with each type of partisan.

4. Personality and Partisanship

Researchers have long been interested in the personality origins of political attitudes and behavior (Adorno et al., 1950; Eysenck, 1954; McClosky, 1958). Throughout this article, I define personality traits as "relatively enduring patterns of thoughts, feelings, and behaviors that distinguish individuals from one another" (Roberts & Mroczek, 2008, p. 31) and that are exogenous to their political socialization (McCourt et al., 1999). From this perspective, the focus on personality traits as determinants of partisan attachments offers two distinct advantages: First, despite some developmental changes in dispositional traits during early adulthood, personality traits are relatively stable, which allows for a more generalizable interpretation of their effects on partisanship throughout an individual's life cycle. Second, personality traits temporally precede the development of many political values, attitudes, and behavior, including party attachments. Thus, despite the observational nature of the following analyses, personality traits intuitively are more likely to be a determinant of partisanship rather than vice versa (see Luttig, 2021, for an exception).

Within the large and diverse share of scholarship on the relationship between personality and politics, there are a few select and distinct traits that are featured quite prominently. These traits include Authoritarianism, SDO, the NfC, as well as the Big Five. While most of the prior literature tends to focus on either one or two of these traits at a time, this article examines the impact of all four personality concepts, thereby offering a systematic and comprehensive overview of the relationship between personality and partisanship. Prior scholarship has also focused much more extensively on the personality origins of PPID (Cooper et al., 2013; Gerber et al., 2012; Schoen & Schumann, 2007) and NPID (Webster, 2018). From this perspective, my predictions for Positive and Negative Partisans are most firmly grounded in prior scholarship, while the determinants of Apathetic and Closed Partisans constitute mostly uncharted territory. I thus remain agnostic regarding their personality

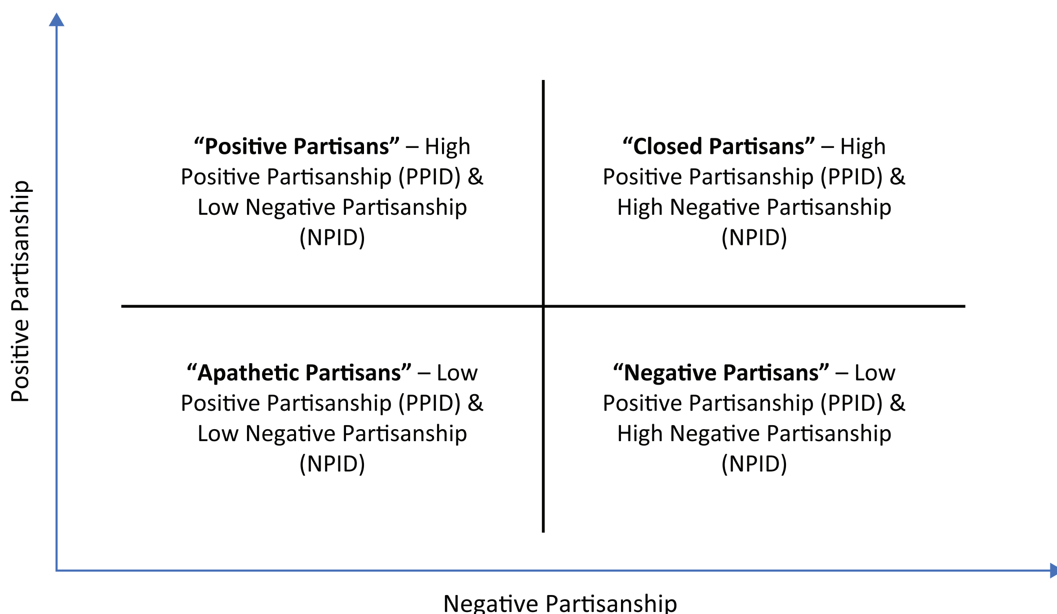

Figure 1. Typology of partisans.

associations. Yet, their exploratory nature also offers new avenues for future research on the origins of PPID and NPID and their varying intensity. In the following, I will briefly elaborate on each trait and articulate my expectations for their effect on PPID and NPID.

4.1. Need for Closure

The NfC is a psychological predisposition that has been used extensively in psychology to describe individuals with a "desire for a firm answer to a question, any firm answer as compared to confusion and/or ambiguity" (Kruglanski, 2004, p. 6). From this perspective, people with high levels of NfC tend to prefer firm and unequivocal assessments of the world and avoid ambiguity and nuance that could negate their need for order and structure (Kruglanski & Webster, 1996).

Accordingly, NfC has been associated with heightened in-party favoritism and out-party hostility as well as partisan identity strength (Luttig, 2018). As Luttig (2018, p. 240) explains:

> Group identification, ingroup bias, and outgroup prejudice are motivated partly by the need for certainty and closure because groups provide members with a social identity and prescribe beliefs about who one is and what they should believe and think. Furthermore…uncertainty as a motivation for group membership can foster extremism, as extreme groups are more distinct and unambiguous.

From this vantage point, NfC might strongly predict high levels of PPID and NPID since they facilitate the rigid categorization of political parties into "good" and "bad," "us" versus "them" (H1a). At the same time, NfC has been linked to political conservatism as well as more right-wing political party preferences (Kossowska & Hiel, 2003), which leads to the expectation that NfC is more strongly related to PPID on the ideological right (H1b).

4.2. Authoritarianism

Authoritarianism reflects a general preference for social conformity over individual autonomy (Feldman, 2003; Feldman & Stenner, 1997; Stenner, 2005), driven by a strong dispositional need for order, certainty, and security as well as a general commitment to conventions and norms (Hetherington & Weiler, 2009; Jost et al., 2003). I thus expect (H2a) that Authoritarianism is positively linked to strong positive partisan identities since they provide a sense of belonging, group norms to comply with, as well as a simplified understanding of who is a friend or foe in a complex political world (see also Luttig, 2017). Negative partisan identities, on the other hand, do not satisfy the need for *inclusion* as easily as positive partisan identities do (for a similar argument, see Zhong, Galinsky, & Unzueta, 2008; Zhong, Phillips, et al., 2008). Instead, NPID turns the *exclusion* from a group—the "not being one of them"—into a meaningful social identity while it provides little affirmational guidance on who we are. Put differently, NPID leaves more uncertainty and imposes less cognitive order than PPID, which is why I expect Authoritarianism to be negatively related to strong NPID (H2b). Like NfC, Authoritarianism is strongly related to ideological conservatism (Federico & Reifen Tagar, 2014) and right-wing policy preferences (Hetherington & Suhay, 2011). Thus, Authoritarianism might be more predictive of PPID on the ideological Right (H2c).

While I treat Authoritarianism as a determinant of PPID and NPID, some prior work has challenged this causal order. In the example of the US, Luttig (2021, p. 786) notes that:

> As the GOP became more conservative on social issues, embraced the religious right, advocated being tough on crime…they communicated that their party sees the world as a dangerous place and that they value obedience, respect, good manners, and good behavior. Inferring the associations of the parties with these values, people change either their psychological worldview or the way that they answer survey questions about these topics to reduce cognitive dissonance.

While it is unclear to what extent this nuance also applies to other personality traits, I acknowledge it and avoid any strict causal claims in the results section.

4.3. Social Dominance Orientation

SDO is another individual-difference variable that is particularly relevant in the study of prejudice. People on the low end of SDO tend to endorse group equality and oppose societal hierarchies, while people on the high end seek power and high status for their group as well as dominance over others (Pratto et al., 1994). SDO thus draws people towards political parties and policies that rationalize and bolster group-based inequalities (Duckitt & Sibley, 2009). Consistent with these expectations, prior research has shown SDO to be strongly related to conservative or right-wing policies and party preferences in the US as well as beyond (e.g., Van Assche et al., 2019). I thus expect that SDO is more predictive of strong PPID on the right both in the US as well as in Sweden (H3a). At the same time, I expect to find a strong connection between SDO and NPID towards the ideological left since many left-wing policies aim to eradicate intergroup inequalities (e.g., affirmative action, access to social services, and universal healthcare) and promote awareness of systemic discrimination and privilege (H3b). It is uncertain how SDO relates to PPID and NPID overall. It is possible that effects cancel each other out once partisans on the left and on the right are jointly examined. I thus remain agnostic regarding their connection to SDO.

4.4. Big Five

The Big Five traits are a well-known and established framework for studying personality which specifies a small set of core traits, including Extraversion, Conscientiousness, Agreeableness, Emotional Stability, and Openness to Experience (see McCrae & Costa, 2008). Prior scholarship has demonstrated the impact of these traits on party preferences on the ideological left and right, albeit with somewhat mixed results. The most consistent finding is the relationship between Openness to Experience and liberalism on the one hand and between Conscientiousness and conservatism on the other (e.g., Alford & Hibbing, 2007; Mondak, 2010; Mondak & Halperin, 2008). There is also some evidence that Emotional Stability is linked to support for conservative candidates and parties and that Agreeableness is connected to support for liberal candidates and parties (Barbaranelli et al., 2007; Caprara et al., 1999; Mondak, 2010). This abundance of prior scholarship provides the foundation for a few concrete hypotheses regarding the relationship between the Big Five and partisan identities on the Left and on the Right. First, and in alignment with prior results, Openness to Experience should be related to a strong PPID on the left (H4a), while Conscientiousness should be related to a strong PPID on the right (H4b). Since Extraversion is connected to social and outgoing behavior, I also expect this trait to predict strong party attachments—regardless of their ideological direction (H4c). Since the evidence on Agreeableness and Emotional Stability is much more mixed, I remain agnostic about their impact.

Prior scholarship is less plentiful regarding the psychological origins of NPID though there is some evidence that Extraversion, Agreeableness, and Emotional Stability are negatively related to strong NPID (Abramowitz & Webster, 2018). Indeed, Extraversion and Agreeableness describe a person who is willing to hear the other side in a polite and trusting manner, while Emotional Stability reduces the chance of experiencing strong negative emotions such as anger and disdain in the first place. Webster (2018) further distinguishes between being a Negative Partisan and the intensity of NPID. The author shows that higher levels of Extraversion are associated with a lower probability of being a Negative Partisan (Webster, 2018). This finding has high face validity since, as Webster (2018) notes, extraverted individuals are more likely to be exposed to a vast array of different political viewpoints. This diverse exposure moderates their negativity towards the out-party and its members. Webster (2018) also demonstrates that higher levels of Agreeableness lessen the degree to which an individual exhibits negative affect toward the out-party and its members since the trait is associated with friendliness, fairness, and decency—even towards the out-party. From this perspective, these three traits—Extraversion, Agreeableness, and Emotional Stability—should be negatively related to strong NPID (H4d). While Abramowitz and Webster (2018) and Webster (2018) use feeling thermometer scales to gauge NPID, I utilize a multi-item scale that measures NPID explicitly as an identity, which I will elaborate on in the next section.

5. Data and Measurements

5.1. US Sample

For the analysis of partisanship in the US, I utilize original survey data that was collected in 2022 by a survey firm called Bovitz Inc., which provides an online panel of approximately one million respondents who participate in multiple surveys over time and receive compensation for their participation. While the sample is not nationally representative, it does reflect the US population on key demographics (see Table A2 in the Supplementary File). The sample includes 1,007 respondents, 882 of them completed the PPID scale, while 876 of them completed the NPID scale. Respondents who identified as a Democrat (or Republican) received the PPID for the Democratic (or Republican) Party and the NPID scale for the Republican (or Democratic) Party. The sample included 456 Republicans and 447 Democrats.

5.2. Sweden Sample

For the analysis of partisanship in Sweden, I utilize original survey data that was collected in 2021 by Bovitz Inc. While the sample is not nationally representative either, it does reflect key demographics of the Swedish population (see Table A3 in the Supplementary File). The sample includes 1,208 Swedish respondents, 968 of them completed the PPID scale, while 975 completed the NPID scale. Respondents received the PPID scale if there was a party that they considered “best” or if they indicated feeling closer to a particular party. Most commonly, that applied to the Social Democrats (28%), the Sweden Democrats (29%), and the Moderate Party (13%). NPID scale was administered based on the question of whether there is a political party that the respondent would never vote for. If so, this party was the target of the NPID scale, which most frequently applied to the Left Party (N = 265), the Green Party (N = 298), the Sweden Democrats (N = 447), and the Feminist Initiative (N = 334). If there were multiple political parties that respondents would never vote for, as is possible in multi-party systems, then respondents were asked to rate these parties on a feeling thermometer scale from 0 to 100. The party with the lowest rating was then selected as the target for the NPID scale.

5.3. Positive Partisanship

I conceptualize and measure PPID as a social identity (see Bankert et al., 2017). The importance of the in-party to an individual’s self-concept, as well as the emotional significance of the membership in that party, is reflected in

items such as "When people criticize this party, it feels like a personal insult" and "When I speak about this party, I usually say 'we' instead of 'they.'" Combining these eight items into one scale yields a continuum that can account for fine gradations in partisan identity strength—an advantageous feature since multi-item partisan identity scales have proven to be more effective than the traditional single-item in predicting political outcomes such as vote choice and political participation in the US and in European (Huddy et al., 2015) multi-party systems (Bankert et al., 2017).

5.4. Negative Partisanship

Social identity theory has also been insightful for the development of a negative partisan identity and its measurement. According to social identity theory, identities cannot only form as a function of common characteristics among in-group members but also in opposition to groups to which we do not belong. Thus, the identity is negative in the sense that it centers on the rejection of the out-group's characteristic. The negative partisan identity scale closely resembles the positive partisan identity scale and captures the emotional significance respondents associate with their rejection of the out-party with items such as "When I meet somebody who supports this party, I feel disconnected" and "I get angry when people praise this party" (see Bankert, 2021).

Descriptive statistics for all key variables and their measurements can be found in Tables A4, A5, and A6 of the Supplementary File, while the distribution of the PPID and NPID scale items can be found in Tables A7 to A10 in the Supplementary File. Pairwise correlations of all key variables are included in Tables A11 and A12.

6. Analyses

In the following analyses, I aim to investigate the personality determinants of PPID and NPID. The correlations between the two in the US are much higher than in Sweden (0.65 versus 0.36), indicating their overlapping nature in the two-party system. From this perspective, simply regressing the personality predictors onto the PPID and NPID values would make it challenging to disentangle the distinct psychological origins of these two types of partisanship. I thus create four different types of partisans based on their values on the NPID and PPID scales. For analytical purposes, "low" is defined as below the sample's mean value on the PPID/NPID scale, while "high" is defined as above the sample's mean value. This strategy preserves sample size while also providing a clear cut-off point.

The percentage shares for each type of partisan are included in Table 1. Both in the US as well as in Sweden, the overwhelming share of partisans fall into the categories of Closed Partisans and Apathetic Partisans; 42% of all American and 43% of Swedish partisans in the sample score highly on both the PPID and NPID scale, while 37% and 24% of American and Swedish partisans respectively are characterized by low scores on both the PPID and NPID scale. Only 10% of American partisans and 16% of Swedish partisans score highly on the PPID scale in conjunction with low values on the NPID scale. Similarly, 11% of American partisans and 17% of Swedish partisans fall on the high end of the NPID scale while also scoring low on the PPID scale. These comparisons reveal an interesting asymmetry: While NPID and PPID can certainly occur independently, the two types of partisanship much more commonly tend to occur together.

In the next part of the analysis, I examine whether these partisan types are related to distinct personality traits (see Figure 2). For this purpose, I regress each dichotomous partisan type onto the personality traits as well as a set of standard control variables (see Table A6 of the Supplementary File). Starting with Positive Partisans in the US, SDO, Agreeableness, and Openness to Experience emerge as strong and positive predictors. The effects of SDO and Openness are quite sizable. Across the range of SDO, PPID increases from 0.04 to 0.17 while keeping all other variables at their mean. There is a similarly steep increase in PPID from 0.02 to 0.1 as Openness increases from 0 to 1. At the same time, NfC is negatively related to being a positive partisan. As NfC increases, the probability of being a positive partisan significantly decreases from 0.12 to 0.03.

Moving on to Negative Partisans, NfC and Emotional Stability emerge as positive predictors of being a Negative Partisan, with similar increases in its predicted probability from 0.05 to 0.17 across the range of these two personality traits. Conscientiousness and Extraversion are uniquely and negatively related to NPID. In combination, these findings suggest that PPID and NPID do have distinct personality origins, in support of the notion that these two are independent constructs.

Table 1. Percentage shares of partisan types.

	US Sample	Swedish Sample
Positive Partisans	10%	16%
Negative Partisans	11%	17%
Closed Partisans	42%	43%
Apathetic Partisans	37%	24%

Notes: Percentages are derived from the sample of respondents who completed both the PPID and NPID scales; N = 1,007 in the US sample and N = 1,208 in the Swedish sample.

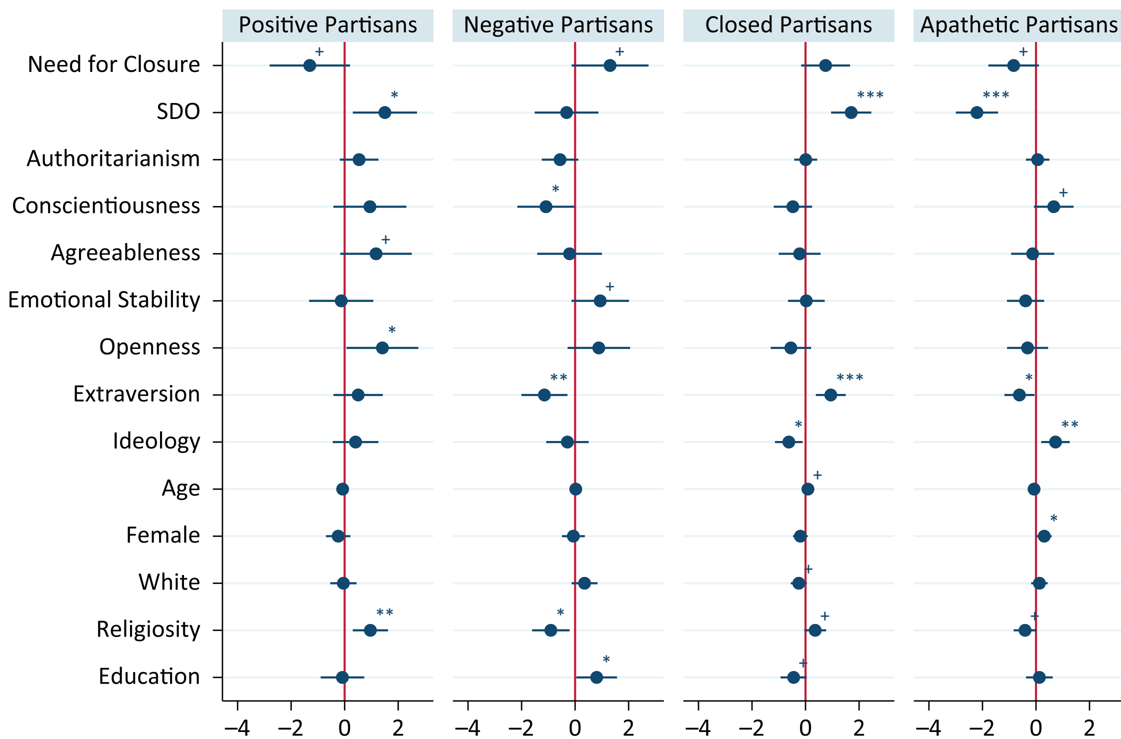

Figure 2. Personality predictors of partisan types: US sample. Notes: Coefficients were estimated using a logistic regression model; all variables are scaled to range from 0 to 1; the corresponding table can be found in Table A13 in the Supplementary File; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Indeed, only NfC appears as a significant predictor in both analyses of Positive and Negative Partisans but with oppositional effects.

Closed Partisans are characterized by high levels of SDO as well as Extraversion. The coefficient for NfC is positive and quite substantial. The predicted probability of being a Closed Partisan increases from 0.26 to 0.67 along the range of SDO—an effect that is similar to Extraversion, which is associated with a growth from 0.28 to 0.51 while holding all other personality variables constant. Last, Apathetic Partisans are characterized by higher levels of Conscientiousness and lower levels of NfC, SDO, and Extraversion. The effects are particularly strong for SDO. Across its range, the predicted probability of being an Apathetic Partisan shrinks from 0.56 to 0.12. Taken together, these analyses suggest that all four types of partisanship have distinct personality profiles. Yet three personality traits—NfC, SDO, and Extraversion—emerge frequently as significant predictors. When adjusted for multiple comparisons using Bonferroni correction, several relationships persist, such as the positive relationship between Negative Partisans and Extraversion, Closed Partisans and SDO and Extraversion, as well as Apathetic Partisans and SDO (see Tables A19 to A22 in the Supplementary File).

These results can only speak to American partisans, which limits their generalizability given the idiosyncratic nature of the US political system. Thus, I replicate the preceding analyses with a sample of Swedish partisans, which illuminates the nature of partisanship in multi-party systems (see Figure 3). For the prediction of Positive Partisans, none of the included personality variables appear to exert an impact which is an interesting departure from the US model. Moving on to Negative Partisans, only Authoritarianism emerges as a negative and significant predictor—which, once again, stands in sharp contrast to the results from the US sample. Indeed, as Authoritarianism increases from 0 to 1, the probability of being a Negative Partisan in Sweden decreases from 0.22 to 0.13.

Among Closed Partisans, SDO (like in the US) and Authoritarianism exert significant effects. As these two traits increase from 0 to 1, Closed Partisanship's likelihood grows from 0.28 to 0.51. Remarkably, these two traits are negative predictors of being an Apathetic Partisan, with a decline in its predicted probability from 0.31 to 0.17 and 0.35 to 0.14 across the range of Authoritarianism and SDO, respectively. Interestingly, NfC and Conscientiousness are positively associated with being an Apathetic Partisan. The positive effects of Conscientiousness, as well as the negative effects of SDO, also surfaced among Apathetic Partisans in the US. These results are robust to alternative model specifications such as a multinomial logistic regression (see Tables A17 and A18 in the Supplementary File). When using Bonferroni-adjusted p-values, the relationship between Closed Partisans and Authoritarianism remains as well as the effect of SDO and Conscientiousness on Apathetic Partisans (see Tables A22 to A26 in the Supplementary File).

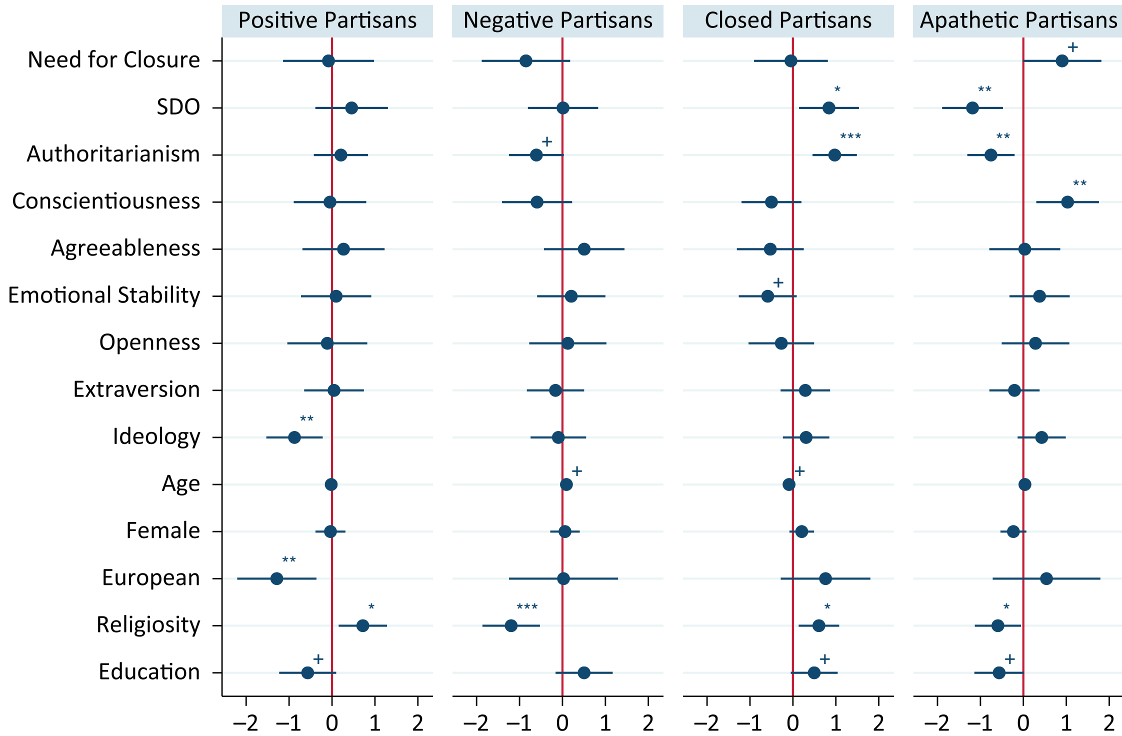

Figure 3. Personality predictors of partisan types: Sweden sample. Notes: Coefficients were estimated using a logistic regression model; all variables are scaled to range from 0 to 1; the corresponding table can be found in Table A14 in the Supplementary Files; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Overall, these results provide two novel insights: First, the four types of partisanship are related to distinct personality profiles both in the US and in Sweden. Second, the four types of partisanship in the US and in Sweden are related to different personality profiles. This variation might speak to the role of institutional features such as the number of political parties (two-party versus multi-party system), the electoral rules (proportional versus majoritarian), the ideological space of the political system, as well as the country's political culture.

7. Ideological Differences Among Positive and Negative Partisans

The preceding analyses revealed distinct personality profiles for each of the four partisan types. Yet it is possible that there are personality differences between Positive and Negative Partisans on the left and right of the ideological spectrum.

To assess this possibility, I first examine the strength of PPID on the right in combination with low levels of NPID towards the left. In the US sample, this involves respondents who identify with the Republican Party but display low levels of NPID towards the Democratic Party. In the Swedish sample, this includes respondents who feel closer to the Moderate Party, the Sweden Democrats, or the Christian Democrats with weak NPID towards the left. Starting with the US (Figure 4), strong positive Republican partisanship is positively related to multiple personality traits, including SDO, Agreeableness, and Extraversion. The significant effects for SDO and Agreeableness remain even when using Bonferroni-adjusted *p*-values (see Table A27 in the Supplementary File). In contrast, Positive Partisans on the right in Sweden (Figure 5) feature *lower* levels of Agreeableness while also, similarly to Republicans in the US, scoring more highly on Extraversion.

For PPID on the left, I examine Democrats in the US. In Sweden, I include respondents who feel closer to the Left Party, the Green Party, the Feminist Initiative, or the Social Democrats. In both cases (see Figures 4 and 5), strong PPID on the left is not related to any personality traits. Only religiosity is a positive determinant in both countries, which is a noteworthy similarity.

Last, I replicate the same analyses for Negative Partisans who disdain certain political parties on the left or the right while being only weakly attached to a political party. In the US sample, this approach includes respondents with NPID towards the Democratic (left) and Republican Party (right), respectively. In Sweden, as exemplars of NPID towards the left, I include respondents who would never vote for the Left Party, the Green Party, the Feminist Initiative, or the Social Democrats. For NPID towards the right, I examine respondents who report never voting for the Moderate Party, the Sweden Democrats, or the Christian Democrats.

In the US (Figure 4), NPID towards the Republican Party is positively related to NfC and Emotional Stability but negatively related to Authoritarianism. In contrast, NPID towards the right in Sweden (Figure 5) is associated

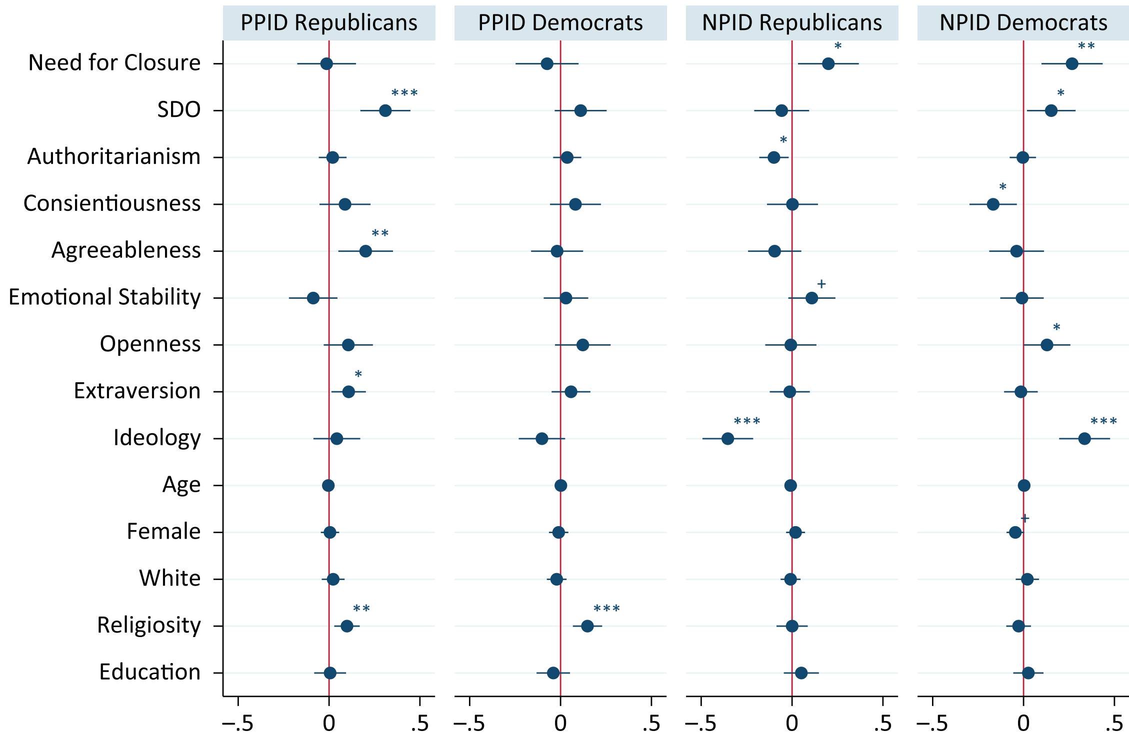

Figure 4. Personality predictors of PPID and NPID among Democrats and Republicans: US sample. Notes: Coefficients were estimated using an OLS regression model; all variables are scaled to range from 0 to 1; the corresponding table can be found in Table A15 in the Supplementary File; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

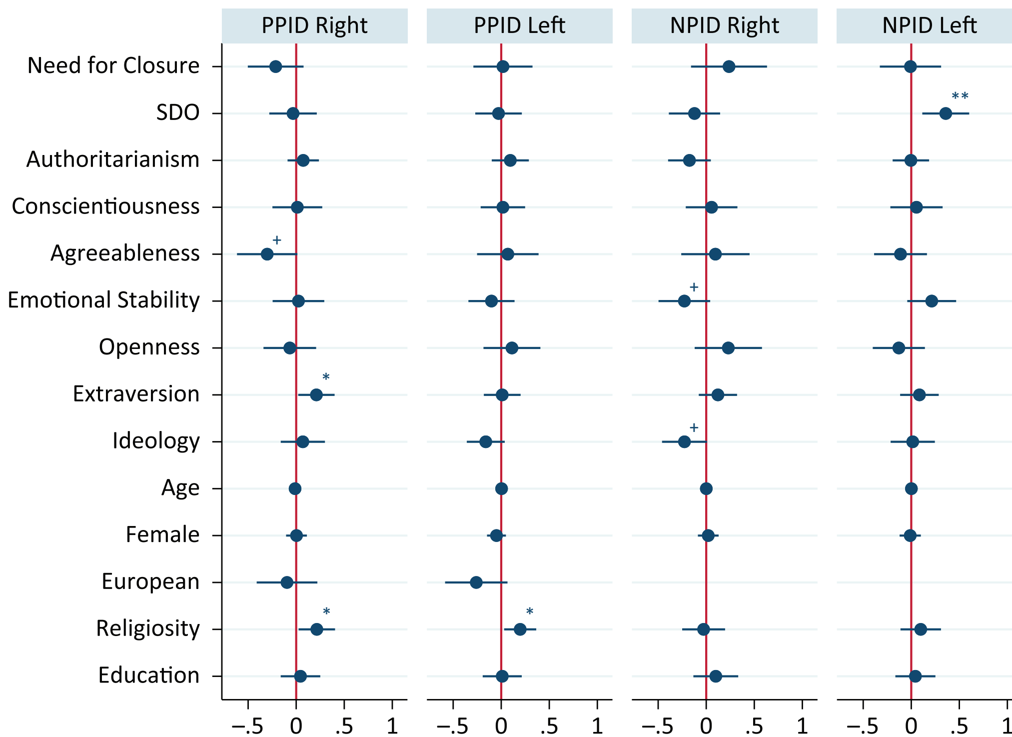

Figure 5. Personality predictors of PPID and NPID among partisans on the left and the right: Sweden sample. Notes: Coefficients were estimated using an OLS regression model; all variables are scaled to range from 0 to 1; the corresponding table can be found in Table A16 in the Supplementary File; the control variable "European" is omitted in the analysis of "NPID Right" and "NPID Left" due to collinearity; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

with *lower* levels of Emotional Stability. Thus, the impact of Emotional Stability is radically different across the US and Sweden. NPID towards the left—i.e., the Democratic Party (Figure 4)—is positively related to NfC, SDO, as well as Openness to Experience, and negatively related to Conscientiousness. The effects are particularly powerful across the range of NfC, whereby strong NPID towards Democrats intensifies from 0.13 to 0.40. The impact of NfC and Conscientiousness is robust to Bonferroni adjustments (see Table A30 in the Supplementary File). In Sweden (Figure 5), NPID towards the left is characterized by high levels of SDO (like in the US). The effects are particularly remarkable for SDO, which is associated with an increase in NPID towards the left from 0.34 to 0.70. This effect also remains when Bonferroni-adjusted *p*-values are used (see Table A34 in the Supplementary File).

Overall, these analyses reveal a few interesting patterns: NfC is positively related to NPID in the US but not in Sweden (H1). Authoritarianism is negatively related to NPID towards the right in the US but not in Sweden (H2). Among Negative Partisans towards a left-wing party, SDO emerges as a common predictor in both the US as well as in Sweden; consistent with H3, a preference for intergroup hierarchies is associated with higher levels of NPID towards the left. Extraversion was positively related to strong PPID on the right in both the US as well as Sweden (partially in alignment with H4c), while Agreeableness was a negative determinant in Sweden but a positive one in the US. This disparity might be reflective of the different political cultures in these two countries, whereby conservatism is much more the *leitkultur* in the US than in Sweden. Last, in both countries, PPID on the left was not associated with any of the included personality traits.

8. Conclusion

This article has examined the personality profiles of four distinct partisan types in both the US and Sweden—two vastly different political systems and cultures. The analyses revealed only a few similarities, such as the role of Extraversion among Positive Partisans on the right as well as the relationship between SDO and NPID towards the left. Overall, however, personality profiles differ across partisan types and across countries. These dissimilarities provide two important insights: First, PPID and NPID are two separate entities that can operate together but also independently of each other. Second, personality predispositions naturally interact with their environment, which might explain the inter-country differences in partisan-personality associations. Indeed, there are important nuances to consider. For example, in a multi-party system that does not foster an "us versus them" mindset, a different type of personality is required to develop NPID in the first place. In other words, the bar might be higher for Swedish partisans to acquire NPID than for their American counterparts who have an instinctive out-party within their two-party system. This also has methodological implications. In this study, American partisans automatically received the NPID scale for the opposition party of their in-party, while Swedish partisans received the NPID scale only if they identified a political party that they would never vote for. This extra step might weaken the comparability of the NPID scales across samples since there might be American partisans who identify with one political party but might still be open to voting for the other. Future research might examine these contextual variations and their implications for measuring PPID and NPID in a comparative setting.

The partisan typology also has implications for political behavior. As the survey data shows in both samples, Positive Partisans are significantly more likely to vote than Negative Partisans, while Closed Partisans are significantly more likely to vote than Apathetic Partisans. At the same time, Closed Partisans are also significantly more likely to agree that "violence might sometimes be necessary to fight against parties and candidates that are bad for this country" and to believe that their "party's opponents are not just worse for politics—They are downright evil." From this perspective, it is especially vital to recognize these different types of partisans and to understand their different psychological origins.

Finally, this manuscript also sheds light on the scope of each partisan type in the electorate; only about 40% of partisans in both samples constitute Closed Partisans, while less than 20% comprise Negative Partisans. That still leaves about 40% of all partisans who are either Apathetic or purely Positive. This should spur more research into how to expand the share of Positive Partisans by, for example, turning Apathetic Partisans into Positive ones.

Acknowledgments

I would like to thank my colleague Stephen Nicholson, Philip H. Alston Jr. Distinguished Professor at the University of Georgia, for his generous support of my research, including the data collection for this article. I would also like to thank Andrew L. Stone, PhD student at the University of Georgia, for prompt research assistance.

Conflict of Interests

The author declares no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Abramowitz, A. I., & Webster, S. W. (2016). The rise of negative partisanship and the nationalization of U.S.

elections in the 21st century. *Electoral Studies*, *41*, 12–22.

Abramowitz, A. I., & Webster, S. W. (2018). Negative partisanship: Why Americans dislike parties but behave like rabid partisans: Negative partisanship and rabid partisans. *Political Psychology*, *39*, 119–135.

Adorno, T., Frenkel-Brenswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The autoritarian personality*. Harper & Brothers.

Alford, J. R., & Hibbing, J. R. (2007). Personal, interpersonal, and political temperaments. *The ANNALS of the American Academy of Political and Social Science*, *614*(1), 196–212.

Bankert, A. (2021). Negative and positive partisanship in the 2016 U.S. presidential elections. *Political Behavior*, *43*(4), 1467–85.

Bankert, A., Huddy, L., & Rosema, M. (2017). Measuring partisanship as a social identity in multi-party systems. *Political Behavior*, *39*(1), 103–132.

Barbaranelli, C., Caprara, G. V., Vecchione, M., & Fraley, C. R. (2007). Voters' personality traits in presidential elections. *Personality and Individual Differences*, *42*(7), 1199–1208.

Bosson, J. K., Johnson, A. B., Niederhoffer, K., & Swann, W. B., Jr. (2006). Interpersonal chemistry through negativity: Bonding by sharing negative attitudes about others. *Personal Relationships*, *13*(2), 135–150.

Caprara, G. V., Barbaranelli, C., & Zimbardo, P. G. (1999). Personality profiles and political parties. *Political Psychology*, *20*(1), 175–197.

Caruana, J., McGregor, R. M., & Stephenson, L. (2015). The power of the dark side: Negative partisanship and political behaviour in Canada. *Canadian Journal of Political Science/Revue Canadienne de Science Politique*, *48*(4), 771–789.

Cooper, C. A., Golden, L., & Socha, A. (2013). The big five personality factors and mass politics. *Journal of Applied Social Psychology*, *43*(1), 68–82.

Duckitt, J., & Sibley, C. G. (2009). A dual-process motivational model of ideology, politics, and prejudice. *Psychological Inquiry*, *20*(2/3), 98–109.

Elsbach, K. D., & Bhattacharya, C. B. (2001). Defining who you are by what you're not: Organizational disidentification and the National Rifle Association. *Organization Science*, *12*(4), 393–413.

Eysenck, H. J. (1954). *The psychology of politics*. Routledge.

Federico, C. M., & Reifen Tagar, M. (2014). Zeroing in on the right: Education and the partisan expression of authoritarianism in the United States. *Political Behavior*, *36*(3), 581–603.

Feldman, S. (2003). Enforcing social conformity: A theory of authoritarianism. *Political Psychology*, *24*(1), 41–74.

Feldman, S., & Stenner, K. (1997). Perceived threat and authoritarianism. *Political Psychology*, *18*(4), 741–770.

Garzia, D., & Ferreira da Silva, F. (2022). Negativity and political behavior: A theoretical framework for the analysis of negative voting in contemporary democracies. *Political Studies Review*, *20*(2), 282–291.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2012). Personality and the strength and direction of partisan identification. *Political Behavior*, *34*(4), 653–688.

Green, D. P., Palmquist, B., & Schickler, E. (2004). *Partisan hearts and minds: Political parties and the social identities of voters*. Yale University Press.

Greene, S. (1999). Understanding party identification: A social identity approach. *Political Psychology*, *20*(2), 393–403.

Greene, S. (2002). The social-psychological measurement of partisanship. *Political Behavior*, *24*(3), 171–197.

Greene, S. (2004). Social identity theory and party identification. *Social Science Quarterly*, *85*(1), 136–153.

Hetherington, M., & Suhay, E. (2011). Authoritarianism, threat, and Americans' support for the war on terror. *American Journal of Political Science*, *55*(3), 546–560.

Hetherington, M., & Weiler, J. D. (2009). *Authoritarianism and polarization in American politics*. Cambridge University Press.

Huddy, L., Mason, L., & Aarøe, L. (2015). Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review*, *109*(1), 1–17.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375.

Kossowska, M., & Hiel, A. V. (2003). The relationship between need for closure and conservative beliefs in Western and Eastern Europe. *Political Psychology*, *24*(3), 501–518.

Kruglanski, A. W. (2004). *The psychology of closed mindedness*. Psychology Press.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*(2), 263–283.

Luttig, M. D. (2017). Authoritarianism and affective polarization: A new view on the origins of partisan extremism. *Public Opinion Quarterly*, *81*(4), 866–895.

Luttig, M. D. (2018). The "prejudiced personality" and the origins of partisan strength, affective polarization, and partisan sorting. *Political Psychology*, *39*(S1), 239–256.

Luttig, M. D. (2021). Reconsidering the relationship between authoritarianism and republican support in 2016 and beyond. *The Journal of Politics*, *83*(2), 783–787.

Mael, F. A., & Tetrick, L. E. (1992). Identifying organizational identification. *Educational and Psychological Measurement*, *52*(4), 813–824. https://doi.org/10.1177/0013164492052004002

Mayer, S. J. (2017). How negative partisanship affects voting behavior in Europe: Evidence from an analysis of

17 European multi-party systems with proportional voting. *Research & Politics*, *4*(1). https://doi.org/10.1177/2053168016686636

McClosky, H. (1958). Conservatism and personality. *American Political Science Review*, *52*(1), 27–45.

McCourt, K., Bouchard, T. J., Lykken, D. T., Tellegen, A., & Keyes, M. (1999). Authoritarianism revisited: Genetic and environmental influences examined in twins reared apart and together. *Personality and Individual Differences*, *27*(5), 985–1014.

McCrae, R. R., & Costa, P. T., Jr. (2008). The five-factor theory of personality. In O. P. John & R. W. Robins (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 159–181). The Guilford Press.

Medeiros, M., & Noël, A. (2014). The forgotten side of partisanship: Negative party identification in four Anglo-American democracies. *Comparative Political Studies*, *47*(7), 1022–1046.

Meléndez, C., & Kaltwasser, C. R. (2021). Negative partisanship towards the populist radical right and democratic resilience in Western Europe. *Democratization*, *28*(5), 949–969.

Mondak, J. J. (2010). *Personality and the foundations of political behavior*. Cambridge University Press.

Mondak, J. J., & Halperin, K. D. (2008). A framework for the study of personality and political behaviour. *British Journal of Political Science*, *38*(2), 335–362.

Mudde, C., & Kaltwasser, C. R. (2018). Studying populism in comparative perspective: Reflections on the contemporary and future research agenda. *Comparative Political Studies*, *51*(13), 1667–1693.

Oscarsson, H., Bergman, T., Bergström, A., & Hellström, J. (2021). *SNS Democracy Council report 2021: Polarization in Sweden*. SNS. https://snsse.cdn.triggerfish.cloud/uploads/2021/03/sns-democracy-council-report-2021-polarization-in-sweden-english-summary.pdf

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763.

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*(1), 31–35.

Rose, R., & Mishler, W. (1998). Negative and positive party identification in post-communist countries. *Electoral Studies*, *17*(2), 217–234.

Samuels, D. J., & Zucco, C. (2018). *Partisans, antipartisans, and nonpartisans: Voting behavior in Brazil*. Cambridge University Press.

Schoen, H., & Schumann, S. (2007). Personality traits, partisan attitudes, and voting behavior: Evidence from Germany. *Political Psychology*, *28*(4), 471–498.

Stenner, K. (2005). *The authoritarian dynamic*. Cambridge University Press.

Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks-Cole.

Tsatsanis, E., Teperoglou, E., & Seriatos, A. (2020). Two-partyism reloaded: Polarisation, negative partisanship, and the return of the left-right divide in the Greek elections of 2019. *South European Society and Politics*, *25*(3/4), 503–532.

Van Assche, J., Dhont, K., & Pettigrew, T. F. (2019). The social-psychological bases of far-right support in Europe and the United States. *Journal of Community & Applied Social Psychology*, *29*(5), 385–401.

Webster, S. W. (2018). It's personal: The big five personality traits and negative partisan affect in polarized US politics. *American Behavioral Scientist*, *62*(1), 127–145.

Zhong, C., Galinsky, A. D., & Unzueta, M. M. (2008). Negational racial identity and presidential voting preferences. *Journal of Experimental Social Psychology*, *44*(6), 1563–1566.

Zhong, C., Phillips, K. W., Leonardelli, G. J., & Galinsky, A. D. (2008). Negational categorization and intergroup behavior. *Personality and Social Psychology Bulletin*, *34*(6), 793–806.

About the Author

Alexa Bankert is an assistant professor in the Department of Political Science at the University of Georgia. Her research examines the origins, measurement, and effects of positive and negative partisanship in the US two-party system as well as in European multi-party systems. Her work has appeared in peer-reviewed journals such as the *Journal of Politics*, *Political Behavior*, and *Electoral Studies*. She is the recipient of several awards, including the Distinguished Junior Scholars Award given by the Political Psychology Section of the American Political Science Association.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 311–324
https://doi.org/10.17645/pag.v10i4.5665

Article

Is Protest Only Negative? Examining the Effect of Emotions and Affective Polarization on Protest Behaviour

Luca Bettarelli *, Caroline Close, and Emilie van Haute

Centre d'Étude de la Vie Politique (CEVIPOL), Université Libre de Bruxelles, Belgium

* Corresponding author (luca.bettarelli@ulb.be)

Submitted: 26 April 2022 | Accepted: 25 September 2022 | Published: 30 December 2022

Abstract
This contribution sheds light on the link between affect and protest behaviors. Using data from a voter survey conducted around the 2019 elections in Belgium, we examine two dimensions of affect: a vertical one, i.e., negative and positive emotions towards politics in general, and a horizontal one, i.e., affective polarization towards fellow citizens. Our findings make three important contributions. First, we identify five distinct classes of respondents depending on their emotions towards politics (apathetic, angry, hopeful, highly emotional, and average). Second, we demonstrate that the combination of both anger and hope is more strongly associated with protest action than anger alone. By contrast, apathy, characterized by an absence of emotions towards politics, is negatively related to protest behavior. Third, we show that affective polarization is a key driver of protest behavior per se. We also show that the two dimensions of affect have distinctive effects. Yet they interact: Affective polarization towards political opponents compensates for the absence of emotions towards politics in general.

Keywords
affective polarization; Belgium; emotions; protest

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Contemporary politics has been increasingly characterized by its affective character (Webster & Albertson, 2022). At the same time, conventional forms of political participation are losing ground, and social uprisings which challenge the established political order are on the rise. In this context, better understanding the affective drivers of protest behaviors appears crucial. This article explores the connection between affect and protest participation.

More specifically, we examine the role of two dimensions of affect on individual protest behaviors: specific, discrete emotions towards politics in general (vertical dimension) and affective polarization towards other party supporters (horizontal dimension). On the one hand, we consider the combined role of specific, discrete negative and positive emotions towards politics in general, tapping into citizens' emotions towards elites and institutions (Capelos & Demertzis, 2018; Vasilopoulou & Wagner, 2017). On the other hand, we look at the horizontal dimension of affect by investigating the role of affective polarization, that is, the tendency among party supporters (the in-party group) to increasingly dislike or resent supporters of other parties (the out-party group), tapping into citizen's feelings towards other fellow citizens (Iyengar et al., 2019; Wagner, 2021; Ward & Tavits, 2019). In addition, we further investigate the

simultaneous effect of emotions and affective polarization and their interactions. We test these expectations by looking at the case of Belgium, using the 2019 RepResent Panel Voter Survey.

Theoretically, we bridge the literature from social movement studies that look at the role of emotions and affective group ties to the process of identity building and collective protest action (Jasper, 1998; Melucci, 1995, p. 45; Polletta & Jasper, 2001), to individual-level research from social and political psychology that investigates the influence of discrete emotions on how citizens process information, evaluate politics, and shape their political preferences and their decision to take part in political processes, in both the electoral and non-electoral arenas (Altomonte et al., 2019; Close & van Haute, 2020; Marcus, 2000). We also go one step further by not only considering the effect of discrete emotions separately but also how the *combination* of various emotions can affect individual protest behaviors.

Our findings make three important empirical contributions. First, at the descriptive level, a latent class analysis (LCA) shows that respondents display different "clusters" of emotions, and we identify five classes of respondents depending on their emotions towards politics: apathetic, angry, hopeful, highly emotional, and average. Second, we show that the vertical and horizontal dimensions of affect are distinctly related to protest behaviors. On the vertical dimension, we demonstrate that the combination of both anger and hope is more strongly associated with protest action than anger alone. By contrast, apathy, characterized by an absence of emotions towards politics, is negatively related to protest behaviors. On the horizontal dimension, we show that affective polarization is significantly related to protest behaviors. We also demonstrate that the two dimensions of affect interact with each other, with high levels of affective polarization compensating for the lack of emotion towards politics, thus pushing apathetic individuals to participate in protest behaviors.

2. Negative Affect and Protest

2.1. Emotions and Protest

We first investigate the vertical dimension of affect on protest behavior. Political psychology has examined the interplay between discrete emotions and individual protest behaviors (Marcus, 2000), such as signing a petition, demonstrating, boycotting (Capelos & Demertzis, 2018), or voting for protest parties (Altomonte et al., 2019; Marcus et al., 2019; Vasilopoulos et al., 2019). Anger was pinpointed as a crucial driver of protest action (Gaffney et al., 2018; Salmela & von Scheve, 2017; Vasilopoulos et al., 2019), as it closely relates to feelings of frustration, indignation (Jasper, 2014b), or ressentiment (Capelos & Demertzis, 2018). By contrast, studies emphasized that fear and anxiety deter individuals from engaging in protest, particularly in autocratic contexts, where the risk of repression and violence is high (Dornschneider, 2020; Nikolayenko, 2022). In democratic contexts, Capelos and Demertzis (2018) show that, during periods of crisis in Greece, anxious people reported a low political activity while those who were angry reported a high degree of participation, especially in violent actions. Looking at voting behavior in the Brexit referendum, Vasilopoulou and Wagner (2017) show that, while anger was positively associated with support for the leave option, fear prompted greater moderation. According to the appraisal-tendency framework (Lerner & Keltner, 2001), fear would enhance individuals' reliance on the evaluation of the situation and would trigger pessimistic risk estimates and risk-averse choices (Valentino et al., 2008), whereas anger would trigger optimistic risk estimates and risk-seeking choices (Lerner & Keltner, 2001). Individual protest behavior is also associated with positive emotions. Capelos and Demertzis (2018) again show that during periods of crisis in Greece, not only angry but also hopeful people reported a high level of engagement in legal and illegal actions. Their findings echo those of Lerner and Keltner (2001), who show that discrete emotions having a dissimilar valence (positive vs. negative), such as anger and hope, or anger and happiness, can lead to similar risk appraisal, i.e., optimistic risk appraisal. Hence, they show that both anger and hope can be associated with goal-oriented behavior. By contrast, some emotions sharing a similar valence, such as anger and fear, can lead to opposite risk appraisals—then, fear and anger would have opposite effects on protest action.

Yet few of these studies look at the combination or simultaneous effect of positive and negative emotions (for exceptions, see Dornschneider, 2020; Landmann & Rohmann, 2020; Nikolayenko, 2022). One has to look at social movement theories to find studies dealing with sets of emotions as crucial elements in the process of collective identity building and as potential drivers of collective action (Jasper, 1998; Polletta & Jasper, 2001). Jasper (2014a, p. 211) refers to protest as being the result of "pairs of positive and negative emotions," such as outrage and hope (Castells, 2012); or as the result of sequences of emotions, such as shame turning into pride through anger in groups sharing a stigmatized identity (Britt & Heise, 2000). This literature points to the role of sets of emotions in creating, nurturing, and potentially breaking a collective movement.

Taking stock of this research, we test how different types of emotional clusters that respondents disclose relate to their level of protest participation. Among the range of emotions, we focus on anger and hope, as they were shown as central factors for mobilization in previous studies. We argue that it is the combination of anger and hope that is the most likely to prompt participation. In other words, being only angry would be less powerful than being angry and hopeful. Hope—the belief that things may change—is also necessary. The underlying mechanism is that because anger and hope have similar

appraisal themes (Lerner & Keltner, 2001), their effects on behavior reinforce each other. Note that we distance ourselves from most studies that look at respondents' emotions towards a specific event. Rather, we measure respondents' level of emotion when they think about politics in general, which is connected to the concept of political resentment vis-à-vis the political elites and institutions (Capelos & Demertzis, 2018). Consequently, we expect that:

> H1: Respondents displaying a combination of high hope and high anger ("highly emotional" respondents) will report a higher level of protest participation, while respondents displaying low levels of both hope and anger (apathy) will report a lower level of protest participation.

2.2. Affective Polarization and Protest

Next to the vertical dimension of emotions towards politics, we focus on a second, horizontal dimension of affect: affective polarization. Initially introduced by Iyengar et al. (2012), affective polarization refers to the tendency among party supporters (the in-party group) to increasingly dislike or resent supporters of other parties (the out-party group). The fast-growing literature has mainly focused on measuring, assessing, and explaining levels of affective polarization across democracies and over time (see, among others, Bettarelli et al., 2022; Bettarelli & Van Haute, 2022b; Druckman & Levendusky, 2019; Gidron et al., 2020; Iyengar et al., 2019; Reiljan, 2020; Wagner, 2021).

Much less is known about the consequences of affective polarization. Iyengar et al. (2019) summarize congruent findings that show that it has negative non-political consequences, as it damages social relations and negatively affects economic behaviors. However, the evidence is more mixed regarding political consequences. Ward and Tavits (2019) demonstrated that higher levels of affective polarization create biases in the perception of party competition, with voters viewing other parties as more extreme. Furthermore, Hetherington and Rudolph (2015) emphasized that it decreases trust among voters. Affective polarization is also associated with resistance to compromise, intolerance, and advancement of their own group over the collective good (Mason, 2018). More worryingly, Kingzette et al. (2021) show how affective polarization in the US undermines support for democratic norms. On the other hand, there is also evidence of the mobilizing power of affective polarization. Ward and Tavits (2019) showed that it enhances the perception that politics has high stakes and that electoral outcomes and success are highly important. Consequently, they show that high levels of affective polarization also lead voters to perceive that participation is crucial, and to higher levels of turnout (see also Harteveld & Wagner, 2022; Wagner, 2021). Others have shown similar dynamics for ideological polarization, which is associated with higher levels of political interest, political information, and electoral participation (Abramowitz & Saunders, 2008; Dalton, 2008). However, these studies focus on electoral participation. What remains unexplored is whether these findings also apply to other forms of participation and if affective polarization has the same mobilizing power on non-institutional participation, especially protest. We expect that it is the case, and we put forward two types of explanations. First, as suggested above, affective polarization is connected to a sense that "something is at stake" and that participating is important. Second, in line with the affective approach to social movements, affective polarization could involve negative affect (fear, hate, anger, outrage) towards political opponents or other societal groups, which, when shared within the group, can have a mobilizing effect (Jasper, 2014a, p. 209). At the same time, positive affect towards other group members (love, compassion, respect, pride) can help create solidarity, keep the group together, and promote participation. We formulate the following hypothesis:

> H2: Respondents displaying high levels of affective polarization will report a higher level of protest participation, while respondents displaying low levels of affective polarization will report a lower level of protest participation.

In the analysis, we will also consider the combined and interactive effect of the horizontal and vertical dimensions of affect, given that the interaction between these two dimensions remains largely unexplored.

3. Data and Methods

3.1. Case Selection

In this article, we focus on Belgium as negative affect and its political consequences remain understudied in this setting (with a few exceptions, see Bettarelli & Van Haute, 2022a, 2022b; Close & van Haute, 2020; van Erkel & Turkenburg, 2020). It is surprising, as Belgium is an ideal case to better understand the role of emotions and affective polarization in multiparty settings.

Belgium is a highly fragmented multiparty system. Since the split of traditional party families along the French–Dutch linguistic divide, Belgium is characterized by two-party systems operating separately (Table 1): Flemish parties compete in Flanders (north of the country), whereas Francophone parties compete in Wallonia (south of the country). We exclude Brussels from our analysis due to its complexity (parties from the two language groups compete on its territory) and data availability (we do not have data about affective polarization for respondents from Brussels).

Furthermore, the relationship dynamics between parties have changed over the last decades. Belgium has long been labeled as a typical consociational democracy

Table 1. List of parties with representation in the federal parliament, 2014–2019 and 2019–present.

Party Family	Flanders	Wallonia
Christian Democrats	CD&V	CDH
Greens	Groen	Ecolo
Regionalists	N-VA	DéFI
Liberals	OpenVLD	MR
Social Democrats	sp.a	PS
Radical Right	VB	PP
Radical Left	PVDA	PTB

with deep social divisions mediated by consensus at the elite level. However, the capacity of the elite of the two main linguistic groups (French and Dutch speakers) to reach agreements has been challenged in recent years, as indicated by the length of government formation at the federal level (De Winter, 2019). This translated into polarizing trends in the ballot box. The 2019 elections saw substantial shifts in party preferences and the rise of the radical left (PVDA-PTB, 12 seats in the Lower Chamber, +10) and radical right parties (VB, which became the second party in Flanders with 18 seats in the Lower Chamber, +15) and the continuing decline of the center, Christian Democratic parties (CD&V, CDH, DéFI). These trends show how Belgium incarnates the understudied and complex character of polarization in multiparty settings.

3.2. Data

Our main data source is the RepResent Panel Voter Survey 2019, conducted by the Excellence of Science consortium of five research teams at the University of Antwerp, Vrije Universiteit Brussel, KU Leuven, Université Libre de Bruxelles, and UCLouvain. It is a rich and original dataset that includes multiple waves (more details in Pilet et al., 2020; Walgrave et al., 2022). We are interested in the first pre-electoral wave of the survey that was conducted from 5 April to 21 May since this wave included questions on protest participation. A total of 7,617 individuals were interviewed. The survey was conducted using computer-assisted web interviewing questionnaires and was distributed by Kantar TNS to their own online panel. Panel participants were selected using a quota sample based on gender, age, education, and region of residency. The final samples slightly differ from the target population, with an overrepresentation of higher educated respondents and the 45–65 age group. Therefore, when we compute variables using the RepResent dataset, we use weights for age, gender, and education.

3.3. Dependent Variable

To grasp respondents' reported participation in protest actions, we make use of the following question: "There are different ways to improve things in Belgium or to be more politically active. How often did you take part in any of the following actions in the past 12 months?" (1 = *never*, 2 = *seldom*, 3 = *sometimes*, 4 = *often*). Nine types of political action were offered, out of which we focus on four: (a) signing petitions, (b) participating in protest or demonstration, (c) boycotting products, and (d) breaking rules for political reasons. Tables 2 and 3 report descriptive statistics for the above items and the correlation matrix, respectively.

Table 2. Descriptive statistics of items of protest participation.

Variable	Obs	Mean	SD	Min	Max
(a) Petitions	7,539	1.99	0.987	1	4
(b) Protest	7,536	1.486	0.8	1	4
(c) Boycotting	7,539	1.997	1.101	1	4
(d) Breaking rules	7,539	1.383	0.728	1	4

Table 3. Correlations matrix among items of protest participation.

Variables	(a)	(b)	(c)	(d)
(a) Petitions	1.000	—	—	—
(b) Protest	0.515	1.000	—	—
(c) Boycotting	0.532	0.453	1.000	—
(d) Breaking rules	0.396	0.512	0.408	1.000

Operationally, we assemble an additive index that sums the four items (Cronbach's alpha is equal to 0.8) to collapse them into a unique indicator of protest. The resulting variable varies from 4 to 16; the higher the index, the more often respondents engage in protest action.

3.4. Independent Variables

3.4.1. Emotions

Our measure of respondents' emotions towards politics is captured by thermometer ratings. While other measurement strategies exist, such as facial or text/sentiment analysis, or physiological responses (Schumacher et al., 2022), ratings are best suited for large survey designs. Furthermore, it matches our choice of measurement of our second independent variable, affective polarization (see Section 3.4.2). Using thermometer ratings for our two independent variables enhances consistency and comparability, especially since we are interested in the combination of the two. We used the following question: "When you think of Belgian politics in general, to what extent do you feel each of the following emotions?" Respondents were offered eight emotions (anger, bitterness, anxiety, fear, hope, relief, happiness, and satisfaction), and a scale ranging from zero (*not at all*) to 10 (*to a great extent*). As previous research pointed to the crucial role of two emotions, one negative (anger) and one positive (hope), in mobilizing protesters, we focus on these two specific emotions. We computed a similar classification of respondents using all emotions. The number and nature of the groupings are very similar, although the distinction between categories is less clear-cut. Regression results are also highly similar to the ones presented in this analysis. As previous studies showed that fear can be negatively correlated to protest, we ran additional regression models with fear as a discrete emotion. Adding fear did not alter our findings. Note that anger and hope are weakly correlated (−0.2). We make use of the LCA to locate respondents into emotional groups. In such a model, a categorical latent (unobserved) variable is used to identify the probability of each individual belonging to a specific emotional category by means of a generalized structural equation model. We obtain the best fit when our sample is split into five emotional groups (see Figure 1). In light of these results, we define Group 1 as average, when respondents register average scores for both hope and anger; Group 2 as apathy, indicating individuals with low scores in each emotion category; Groups 3 and 4 as hopeful and angry, respectively, where the former includes people with high rates of hope and low rates of anger, while the latter is the other way round; Group 5 as highly emotional, which includes individuals showing high rates of both anger and hope. In the empirical analysis, the average will represent the baseline category.

Other methods than LCA could have been used to assess the combined and isolated effect of discrete emotions, such as interaction effect or principal component analysis (PCA). By using the interaction effect between anger and hope, we would capture the "mediating" effect of one emotion on the other, but we would not actually catch the effect of having both emotions at the same time. Interaction tells us if, e.g., the effect of anger towards politics on protest participation is higher when the degree of hope increases—which is not our argument. Besides, interaction would have implied treating the two emotions, not as continuous but as discrete variables, and then checking all possible combinations of anger and hope. This latter exercise makes the presentation of results much more complex. In the end, the conclusions are highly similar to those when using LCA.

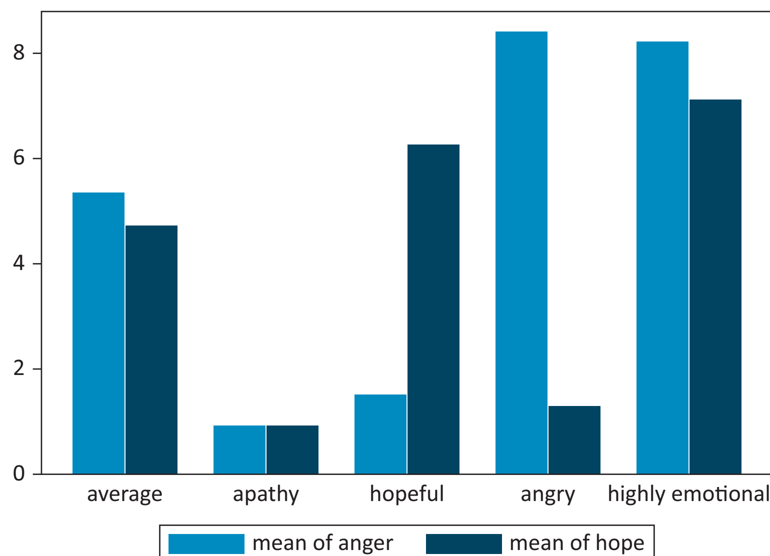

Figure 1. Distribution of anger and hope across groups (LCA).

We also ran a PCA that highlighted two components: one that negatively correlates anger and hope and one that positively correlates the two, each explaining approximately half of the variation in the data. Each component explains half of the scenarios resulting from the LCA: The component that positively correlates the two emotions would contrast the "highly emotional" scenario with the "apathy" one. As a result, we should have used both in the same regression, making results qualitatively similar to LCA but more complex to interpret.

We have explored these possibilities, but LCA offers the best tool to explore our research question and test our hypothesis, both in terms of conceptual message and clarity of presentation of results. LCA groups observations based on a data-driven process. As a result, it creates clear-cut categories that classify respondents based on their emotional states and allows them to then link them to participation in protest action. Conceptually, categorization by means of LCA allows one to clearly disentangle the propensity to participate in protests of different categories of individuals, defined according to their political emotions, with a particular focus on highly emotional people, i.e., those who display high levels of both anger and hope.

In terms of size (Table 4), two groups (average and negative) account for over 70% of the respondents. Nevertheless, no group contains less than 500 individuals. Note that the overall standard deviation of each emotion is consistently larger than that within each group, thus further supporting our modeling choice.

Table 5 reports the distribution of protest participation by group. It indicates that protest participation is significantly lower in the apathy group, and larger in the negative and (mostly) the highly emotional groups, compared to the average. By contrast, no significant differences emerge between the average and positive groups.

3.4.2. Affective Polarization

Contrary to our measurement of discrete emotions, affective polarization does not take into account different feelings towards politics (political elites and

Table 4. Descriptive statistics of anger and hope across groups.

	N	Mean	SD	Min	Max
Overall					
Anger	7,471	5.94	2.69	0	10
Hope	7,469	3.84	2.56	0	10
Average					
Anger	3,408	5.35	1.31	2	9
Hope	3,409	4.99	1.18	3	8
Apathy					
Anger	523	0.93	1.14	0	4
Hope	523	0.91	1.10	0	3
Hopeful					
Anger	612	1.50	1.14	0	4
Hope	612	6.25	1.10	4	10
Angry					
Anger	2,019	8.41	1.43	4	10
Hope	2,017	1.28	1.36	0	3
Highly Emotional					
Anger	699	8.22	1.19	4	10
Hope	699	7.12	1.58	6	10

Table 5. Descriptive statistics of protest participation split by groups.

Categories	N	Mean	SD	Min	Max	p(x,y)
Average	3,405	6.68	2.69	4	16	—
Apathy	521	5.86	2.41	4	15	0.00
Hopeful	610	6.74	2.54	4	16	0.76
Angry	2,093	6.95	2.82	4	16	0.03
Highly emotional	694	8.05	3.41	4	16	0.00

Note: p(x,y) in the last column is the t-test of equality of means across the baseline category *average* (x) and other categories (y), under the assumption of equal variances.

institutions); it is rather a general measurement of negative affect towards other partisan groups. To measure affective polarization for each respondent, we also use thermometer ratings, the most common strategy in the literature (Iyengar et al., 2019). We make use of the following question from the RepResent dataset: "Could you use the scale below to indicate how you feel about the following groups?" (scale ranging from 0 to 100, 0 to 49 = *not very favorable*; 50 = *neutral*; 51 to 100 = *favorable*). The higher the score, the higher the sympathy towards partisans of the party. Following the distinct party offer, respondents in Flanders had to indicate their feelings towards supporters of the seven Dutch-speaking parties listed in Table 1, and respondents in Wallonia had to do the same with the seven French-speaking parties listed in Table 1. We make use of the spread-of-score method proposed by Wagner (2021). The index is computed based on the following equation:

$$\text{Spread}_i = \sqrt{\sum_{p=1}^{P}\left(\text{like}_{ip} - \overline{\text{like}_i}\right)^2}$$

where subscripts i and p indicate each survey respondent and each French- or Dutch-speaking party, respectively; "like" signifies the like–dislike evaluation towards a party on a scale from 0 to 100; and "$\overline{\text{like}}$" is the average like–dislike score of respondent i. The higher the index, the higher the degree of affective polarization. Note that we do not weigh the index for the electoral size of each party, for two reasons. First, as we use Wave 1 (pre-electoral) of the RepResent dataset, we do not have a good reference time point to weigh each party's size. Second, we argue that the weighting strategy is appropriate when using a territorial approach, as the social consequences of disliking supporters of small or large parties may differ significantly; however, it is not essential for individual-level analyses.

The average level of affective polarization among our population is 19.5, ranging from 0 to 49. In Table 6, we describe the average level of protest participation for different intensities of affective polarization (split in percentiles, from <20th to >80th). Results show that participation in protest action significantly increases across percentiles, thus indicating a positive correlation between affective polarization and protest.

3.4.3. Controls

We include standard individual-level socio-demographic variables (gender, age, education) that contribute to determining political engagement (Brady et al., 1995; Marien et al., 2010). Gender is a dummy equal to one for female. Age ("What is your age?") is a continuous variable, while education is a five-category variable, ranging from *none or elementary* to *university degree*. Income is measured by the following question: "To what extent are you satisfied with your family's total income?" (0–10 scale, with 0 = *very unsatisfied* and 10 = *very satisfied*).

We also control for political attitudes. First, we control for respondents' left–right self-placement for two reasons: It allows us to further establish that affective polarization and ideological positions are two distinct phenomena. Second, it allows us to control for the specific dynamics of protest in Belgium under the 2014–2019 legislature. The coalition government that came out of the 2014 elections was exceptional, as it only included (center-)right parties (N-VA, CD&V, OpenVLD, and MR). Therefore, one can expect that protests were initiated by the left-wing opposition. Second, we control for the degree of ideological extremeness of respondents, computed as the difference in absolute value between the score on the left–right scale of each respondent and the average score across our sample. The higher the score; the more ideologically extreme the respondent. Third, we control for respondents' satisfaction with democracy to further establish that emotions towards politics are distinct from evaluations of the political system. We use the following question: "In politics, people often talk of 'left' or 'right.' Can you place your own opinions on a scale from 0 to 10, with 0 meaning *left*, 5 the *centre*, and 10 the *right*?" "Overall, how satisfied are you with the way democracy is working in Belgium?" (1 = *very satisfied*; 5 = *very unsatisfied*). Fourth, we include a variable measuring the respondents' degree of interest in politics on a scale of 0–10, ranging from 0 = *not interested at all* to 10 = *extremely interested*.

Finally, we control for the place of residence of each respondent, as there may exist habits of protest participation linked to territories. To do so, we use NUTS-3 fixed effects in our regression model.

Table 6. Descriptive statistics of protest participation split by percentiles of affective polarization.

	Mean	SD	Min	Max
AffPol < 20	6.36	2.82	4	16
20 < AffPol < 40	6.57	2.63	4	16
40 < AffPol < 60	6.60	2.59	4	16
60 < AffPol < 80	6.88	2.63	4	16
Affpol > 80	7.33	3.01	4	16

Note: The first column indicates the group under analysis with respect to percentiles of the affective polarization distribution, i.e., 20th, 40th, 60th, and 80th.

3.4.4. Modelling Strategy

In our regression analyses, continuous variables were standardized using the z-score, i.e., mean equal to zero and standard deviation equal to one, to ease the interpretation of coefficients among variables computed at different scales. Coefficients were computed using OLS models, with NUTS-3 fixed effects. The dependent variable in the models is the additive index of protest, as discussed above. We checked the presence of potential collinearity issues using the Variance Inflation Factor test and registered a value below two in all models.

4. Results

Table 7 presents the results of our regression analyses. First, we introduce the groups of respondents by type of emotion (Column 1). Coefficients associated with these groups must be interpreted as differences with respect to the baseline group (average anger and hope). Results provide very interesting and novel insights. Protest participation, as expected, is significantly lower in the apathy group compared to the average category. In fact, a switch from the latter to the former increases protest participation by over one point. When we consider the hopeful group, we see that the coefficient is not statistically significant. This denotes that being hopeful when thinking about politics does not represent a sufficient condition per se to increase participation in protest action. However, as shown in Figure 1, this could also result from the fact that levels of hope do not diverge so much between the average and hopeful groups. Turning towards the angry group, we see a positive coefficient, even if it is not statistically significant at any conventional

Table 7. Regression results.

	(1) Protest	(2) Protest	(3) Protest	(4) Protest
Affpol (std)		0.362*** (0.042)	0.353*** (0.043)	0.085* (0.045)
Groups (Emotions)				
Average (baseline)				
Apathy	−1.142*** (0.136)		−1.04*** (0.146)	−0.528*** (0.139)
Hopeful	−0.128 (0.127)		−.316** (0.133)	−0.319** (0.133)
Angry	0.043 (0.09)		−0.042 (0.094)	0.171* (0.097)
Highly emotional	1.363*** (0.173)		1.016*** (0.175)	0.828*** (0.153)
Gender				−0.13 (0.079)
Age (std)				−0.39*** (0.041)
Education (std)				0.113*** (0.04)
Income (std)				−0.078* (0.042)
Left_right (std)				−0.254*** (0.042)
Extremeness (std)				0.227*** (0.045)
Satisfaction with democracy (std)				−0.044 (0.048)
Political interest (std)				0.836*** (0.045)
Observations	6,894	5,990	5,829	5,753
R^2	0.073	0.049	0.072	0.192
NUTS-3 dummies	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

level. These results indicate that positive or negative emotions alone do not contribute to increasing the propensity to engage in protest action. Protest participation is significantly larger than in the average group only for the highly emotional group, with a value of the coefficient much larger than any other category. In fact, the difference between apathetic and highly emotional respondents, those registering the lowest and highest probability of protest participation, respectively, is over two points. These results provide a relevant message: Protest participation is connected to the joint action of positive (hope) and negative (anger) emotions towards politics, thus supporting H1. In other words, those who participate in protest action most frequently feel angry yet hopeful and may believe that political conditions are likely to improve thanks to collective action. Similar findings were uncovered using other statistical methods (interaction effect between discrete emotions and PCA): (a) Anger per se matters more than hope; (b) the interaction between the two has a positive and significant effect; (c) protest participation is particularly high among individuals that show high values of both hope and anger. In addition, if we used LCA categorization based on all emotions available (listed in Section 3.1.1), again, the findings are similar.

In Columns 3 and 4, we introduce the affective polarization index, without and with the groups of emotions, respectively. We do so to test the stability of the coefficient associated with affective polarization when included simultaneously with other political and emotional states. Results reassure us of the independent relationship between affective polarization, as its coefficient does not vary much between the two models. As indicated by Column 3, all else equal, affective polarization positively correlates with participation in protest action (H2 supported). Note that the coefficient associated with the hopeful group becomes significant (with a negative sign). This indicates that, even if positive feelings towards national politics are broadly related to a lower propensity to protest, there may be a subset of politically hopeful respondents who engage in protest action because they dislike (some of) supporters of other parties, thus partially biasing results in Column 1.

Finally, in Column 4, we test if previous results are robust to the inclusion of the set of controls introduced in Section 3.3.3. Before commenting on our main explanatory variables (i.e., emotional groups and affective polarization), we observe the behavior of the controls. The sign and significance level of coefficients denote that participation in protest action is not linked to the gender of respondents, while it is higher among younger respondents with a higher level of education. Moreover, it is lower among well-off people. If we switch the attention to political-related controls, we note that participation in protest action is also higher among respondents who position themselves to the left of the ideological spectrum, or among those who hold more extreme political views. Finally, the degree of satisfaction with democracy does not report a significant coefficient; contrarily, the degree of interest in politics of respondents turns out to be positively correlated to protest participation, thus indicating that the more the respondents are interested in politics, the higher the frequency with which they participate in protest action.

As expected, both the magnitude and significance level of coefficients associated with our variables of interest have changed due to the introduction of the control variables. However, the overall message we can draw from the analysis remains qualitatively similar. As far as the emotional groups are concerned, we can observe that protest participation is significantly larger in the highly emotional group than in any other category, although the size of the coefficient is partially reduced. The apathy group is still characterized by its lower propensity to protest, even if its associated coefficient reduces in magnitude with respect to Column 3. Finally, if the hopeful group behaves consistently with previous findings, the angry one is now significantly correlated with protest participation, even if the magnitude and significance level of the coefficient are somewhat weak. However, what changes the most if compared to previous results is the effect of the affective polarization index, which is now one-fourth of that in Column 3. This is due to the fact that affective polarization captured the effect of some of our controls. In order to fully understand the mechanisms driving this result, we re-estimated the model in Column 3 by adding one control at a time. We do not report the results of this exercise in this article for the sake of brevity, but they are available upon request. We find that only the degree of extremeness and interest in politics affect our findings regarding affective polarization. When we include the degree of extremeness, the coefficient of affective polarization reduces from 0.353 (Column 3) to 0.228 (p-value 0.00). This (partial) reduction can be explained by the fact that extreme voters may have more extreme (negative) feelings towards supporters of other parties. However, as shown in Bettarelli and Van Haute (2022a), affective polarization also operates from moderate to extreme voters, thus leaving room for an independent effect of the affective polarization index. Contrarily, when Polint is included, the coefficient of affective polarization drops from 0.353 (column 3) to 0.102 (p-value 0.18), signaling on one side that people who show high levels of affective polarization are those who care the most about politics and, on the other side, that affective polarization is not only related to affect but is also greatly connected to cognitive processes.

In addition to these analyses, we want to explore further how the two dimensions of affect relate to protest participation. We are interested in how the two dimensions interact, as their combined impact on protest participation may vary according to the specific combination of the two. The interaction between these two dimensions of affect remains largely unexplored, yet it could provide novel insight concerning the drivers of protest participation. For instance, affective polarization may

operate either as a substitute or a complement of emotions. In the former scenario, a high degree of affective polarization would compensate for the lack of emotion towards politics. Or, in other words, horizontal affects driving people to participate in protest action predominate over vertical. Contrarily, if the effect of affective polarization is stronger when emotions are high, it would signal that the horizontal and vertical dimensions reinforce each other.

To better investigate this interaction and to ease the interpretation of results, we collapse the categories of emotions into a unique continuous variable by means of a PCA involving individuals' self-reported degree of anger and hope. We consider the component that positively correlates the two emotions. The resulting variable, which we refer to as political feelings (pol_feel), ranges from ca. −2.6 to +2.7, with higher values corresponding to higher degrees of both hope and anger (i.e., the highly emotional category). Figure 2 below shows the mean of pol_feel by groups of emotions and further corroborates the validity of the PCA exercise.

Next, we run a regression model where we interact the two variables of interest, namely affective polarization and pol_feel, together with the standard set of controls as in Column 4 (Table 7). Figure 3 below plots the average marginal effects. Results suggest a substitution dynamic: The effect of affective polarization is large in magnitude and statistically significant for low levels of pol_feel (i.e., apathetic respondents). Contrarily, its impact drastically decreases as much as the pol_feel index increases, and it becomes not significant for the highly emotional respondents.

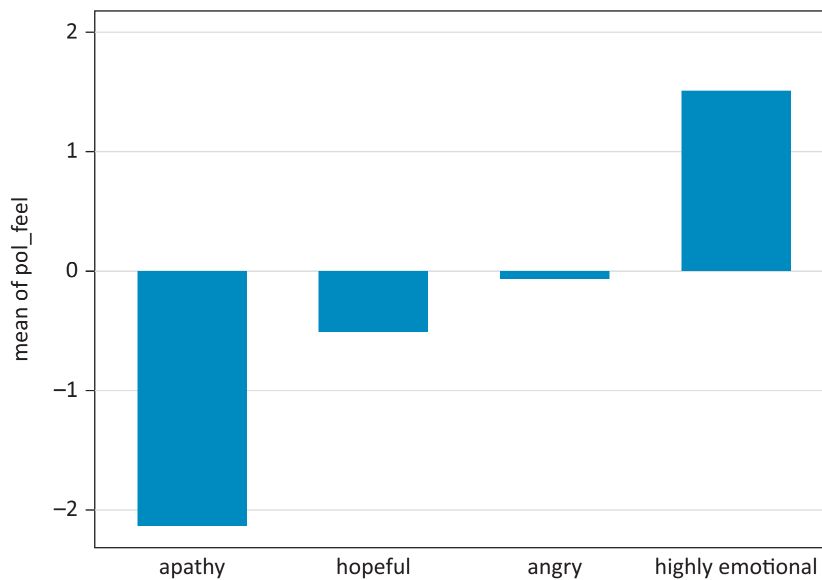

Figure 2. Mean of pol_feel, by groups.

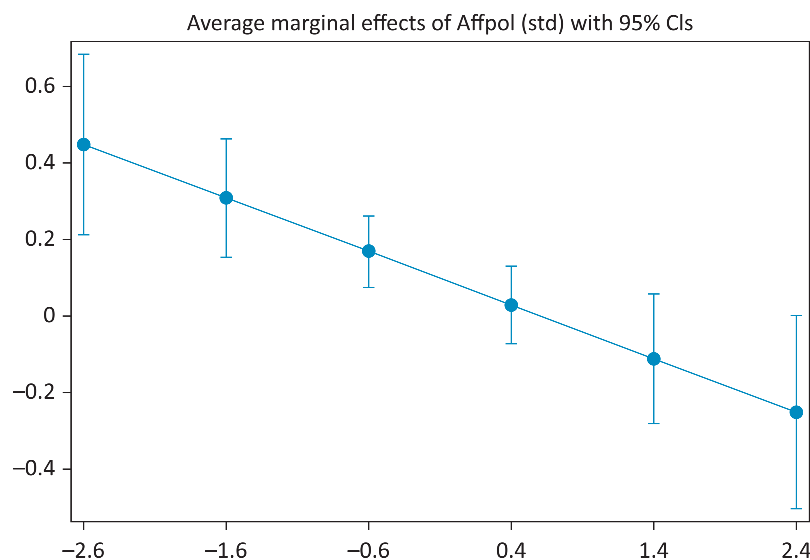

Figure 3. Average marginal effect of affective polarization on protest participation when pol_feel increases (i.e., from no emotions to high emotions). Note: CI stands for Confidence Interval.

To sum up, affective polarization is positively related to protest participation. Moreover, it acts as a substitute for apathy and may stimulate protest participation among people who do not have strong emotions towards politics.

5. Discussion and Limitations

Our analysis has focused on the specific case of Belgium in the 2019 (pre-)electoral sequence. This raises the question of the generalisability of the findings.

Belgium is often described as a multipartisan consociational democracy characterized by a culture of political compromise and by social concertation (Delwit, 2022; Deschouwer, 2012). It has consequences on action repertoires. Protest action, especially in the form of mass demonstrations, is quite common and structured by civil society organizations such as unions. These organizations are linked to the state, which is relatively permeable and open to social movements. Demonstrations are usually peaceful and welcome a broad range of citizens—not exclusively the most extremist or desperate activists. In majoritarian democracies where social movements and the state operate in a more confrontational relationship (such as France or the US), one could expect a stronger and more isolated effect of anger and a lower effect of hope, as well as a stronger effect of affective polarization. Yet while the consociational nature of the Belgian political system usually produces broad coalition governments representing most segments of society, the 2014–2019 government leaned particularly towards the right end of the spectrum, affecting the capacity of left-wing movements (and especially workers' organizations) to influence government policies. This created a climate of social unrest. The relationship between left-wing self-placement and protest uncovered in the models partially reflects that context. Moreover, this context may have exacerbated the role of negative feelings between social groups, hence, affective polarization, as well as distrust and dissatisfaction towards the national government. This specific context tends to offset the specificities mentioned above and brings our findings closer to what one could expect in majoritarian democracies, enhancing the generalizability of our results to other settings.

In addition, 2019 was particularly marked by climate mobilizations, including school strikes and demonstrations (Wouters et al., 2022). Given the nature and objectives of these pro-environmental collective actions, our findings may overestimate the role of positive emotions such as hope (Landmann & Rohmann, 2020). However, our analyses focus on the general population and not the specific segment of climate activists. Wouters et al. (2022) also show that participants in climate mobilizations were younger and less politically experienced than typical demonstrators. This could partly explain the negative relationship we uncover between age and protest, although it is a common pattern in protest participation (Marien et al., 2010).

Our findings are also limited by the methodology adopted in the study. Given that all our measurements are from the same wave of the RepResent survey, the design prevents us from asserting any causal relationship between the variables, nor can we be sure about the direction of any such potential relationship. While theoretically, we could expect affect to influence political behavior, participation in protest action could also create or reinforce emotions towards politics, both negative and positive, as well as affective polarization. Social interactions with like-minded peers in collective action, for instance, could reinforce affective predispositions, which are *shared* within the group, as suggested by social movement theories. The roots of the emotional reactions investigated in this article would deserve specific attention.

Finally, our study does not examine the mechanisms linking affective states to protest action. Emotions may lead to (negative) evaluation or judgment about politics (see Webster, 2018), and this judgment would lead to action. In this case, emotions would *indirectly* influence protest behavior. But emotions could also derive from a cognitive appraisal of the situation and could work as a catalyst for engagement in protest behavior. Regarding affective polarization, we discussed two mechanisms in our theoretical section: One connects affective polarization to political interest and politicization; another connects affective polarization to in- and out-group identity-building dynamics. Yet our empirical strategy does not allow us to disentangle these underlying mechanisms, and further research is needed to provide greater insight in this regard. Our findings nevertheless contribute to stimulating the debate.

6. Conclusion

This article has sought to better understand the role of affect in protest behaviors. We investigate two dimensions of affect. On the vertical dimension, we go beyond the effect of one discrete emotion at a time. Our LCA distinguishes five categories of citizens based on their emotions towards politics: apathetic, angry, hopeful, highly emotional, and average. This is the first important contribution of the article: We show how sets of positive and negative emotions can combine simultaneously in diverse manners and "produce" types of citizens who respond emotionally to politics in very different ways. The behavioral consequences of these combinations deserve further attention. We demonstrate that protest behaviors are the highest among citizens displaying a combination of high anger and hope, and the lowest among apathetic citizens who display an absence of positive or negative emotions towards politics. This is the second important contribution of the article. We show that being angry can mobilize protesters but that the combination of anger and hope can be even more connected to protest action than anger alone. Hope—suggesting a positive or optimistic appraisal of the future—can be crucial

for political engagement. Yet hope alone does not seem to activate protest.

On the horizontal dimension, we show that protest behaviors are highest among citizens displaying higher levels of affective polarization, that is, higher levels of dislike of political opponents. In this case, negativity is key. Interestingly, we also show that the two dimensions are distinct drivers of protest. This is the third important contribution of the article: Affect is crucial to better understand protest behaviors and different dimensions of affect matter.

Lastly, we show that the two dimensions of affect interact. We knew from previous research that affective polarization has mobilization potential. We now better understand how this mobilization works: By appealing to a different dimension of affect, it can mobilize otherwise apathetic citizens. This is the fourth important contribution of the article. Nurturing a dislike of political opponents can make up for the absence of emotions towards politics. This could be a key to better understanding the dynamics of radical parties and leaders.

Acknowledgments

Bettarelli received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 801505. Van Haute received funding from the FWO-FNRS EoS Grant No. 30431006.

Conflict of Interests

The authors declare no conflict of interests.

References

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, *70*(2), 542–555.

Altomonte, C., Gennaro, G., & Passarelli, F. (2019). *Collective emotions and protest vote* (CESifo Working Paper No. 7463). CESifo.

Bettarelli, L., Reiljan, A., & Van Haute, E. (2022). A regional perspective to the study of affective polarization. *European Journal of Political Research*. Advance online publication. https://doi.org/10.1111/1475-6765.12548

Bettarelli, L., & Van Haute, E. (2022a). Affective polarization and coalition preferences in times of pandemic. *Frontiers in Political Science*, *4*. https://doi.org/10.3389/fpos.2022.945161

Bettarelli, L., & Van Haute, E. (2022b). Regional inequalities as drivers of affective polarization. *Regional Studies, Regional Science*, *9*(1), 549–570. https://doi.org/10.1080/21681376.2022.2117077

Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A resource model of political participation. *American Political Science Review*, *89*(2), 271–294.

Britt, L., & Heise, D. (2000). From shame to pride in identity politics. In S. Stryker, T. J. Owens, & R. W. White (Eds.), *Self, identity, and social movements* (pp. 252–270). University of Minnesota Press.

Capelos, T., & Demertzis, N. (2018). Political action and resentful affectivity in critical times. *Humanity & Society*, *42*(4), 410–433.

Castells, M. (2012). *Networks of outrage and hope: Social movements in the internet age*. Polity Press.

Close, C., & van Haute, E. (2020). Emotions and vote choice: An analysis of the 2019 Belgian elections. *Politics of the Low Countries*, *2*(3), 353–379.

Dalton, R. J. (2008). The quantity and the quality of party systems: Party system polarization, its measurement, and its consequences. *Comparative Political Studies*, *41*(7), 899–920.

Delwit, P. (2022). *La vie politique en Belgique de 1830 à nos jours* [The political life in Belgium from 1830 to nowadays]. Editions de l'Université de Bruxelles.

Deschouwer, K. (2012). *The politics of Belgium: Governing a divided society*. Red Globe Press.

De Winter, L. (2019). Government coalitions as a reflection of national politics: The complex case of Belgium. In M. Evans (Ed.), *Coalition government as a reflection of a nation's politics and society* (pp. 64–85). Routledge.

Dornschneider, S. (2020). *Hot contention, cool abstention: Positive emotions and protest behavior during the Arab Spring*. Oxford University Press.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization. *Public Opinion Quarterly*, *83*(1), 114–122.

Gaffney, A. M., Hackett, J. D., Rast, D. E., III, Hohman, Z. P., & Jaurique, A. (2018). The state of American protest: Shared anger and populism. *Analyses of Social Issues and Public Policy*, *18*(1), 11–33.

Gidron, N., Adams, J., & Horne, W. (2020). *American affective polarization in comparative perspective*. Cambridge University Press.

Harteveld, E., & Wagner, M. (2022). Does affective polarisation increase turnout? Evidence from Germany, the Netherlands and Spain. *West European Politics*. Advance online publication. https://doi.org/10.1080/01402382.2022.2087395

Hetherington, M. J., & Rudolph, T. J. (2015). *Why Washington won't work: Polarization, political trust, and the governing crisis*. University of Chicago Press.

Iyengar, S., Lelkes, Y., Levendusky, M. S., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405–431.

Jasper, J. M. (1998). The emotions of protest: Affective and reactive emotions in and around social movements. *Sociological Forum*, *13*(3), 397–424.

Jasper, J. M. (2014a). Constructing indignation: Anger dynamics in protest movements. *Emotion Review*,

6(3), 208–213.

Jasper, J. M. (2014b). Emotions, sociology and protest. In C. von Scheve & M. Salmella (Eds.), *Collective emotions: Perspectives from psychology, philosophy, and sociology* (pp. 341–355). Oxford University Press.

Kingzette, J., Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M. S., & Ryan, J. (2021). How affective polarization undermines support for democratic norms. *Public Opinion Quarterly*, *85*(2), 663–677.

Landmann, H., & Rohmann, A. (2020). Being moved by protest: Collective efficacy beliefs and injustice appraisals enhance collective action intentions for forest protection via positive and negative emotions. *Journal of Environmental Psychology*, *71*, Article 101491. https://doi.org/10.1016/j.jenvp.2020.101491

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146–159.

Marcus, G. E. (2000). Emotions in politics. *Annual Review of Political Science*, *3*, 221–250.

Marcus, G. E., Valentino, N. A., Vasilopoulos, P., & Foucault, M. (2019). Applying the theory of affective intelligence to support for authoritarian policies and parties. *Political Psychology*, *40*(S1), 109–139.

Marien, S., Hooghe, M., & Quintelier, E. (2010). Inequalities in non-institutionalised forms of political participation: A multi-level analysis of 25 countries. *Political Studies*, *58*(1), 187–213.

Mason, L. (2018). *Uncivil agreement: How politics became our identity*. University of Chicago Press.

Melucci, A. (1995). The process of collective identity. In H. Johnston & B. Klandermans (Eds.), *Social movements and cutlure* (pp. 41–63). Routledge.

Nikolayenko, O. (2022). "I am tired of being afraid": Emotions and protest participation in Belarus. *International Sociology*, *37*(1), 78–96.

Pilet, J.-B., Baudewyns, P., Deschouwer, K., Kern, A., & Lefevere, J. (2020). *Les Belges haussent leur voix* [The Belgians raise their voice]. Presses Universitaires de Louvain.

Polletta, F., & Jasper, J. M. (2001). Collective identity and social movements. *Annual Review of Sociology*, *27*, 283–305.

Reiljan, A. (2020). "Fear and loathing across party lines" (also) in Europe: Affective polarisation in European party systems. *European Journal of Political Research*, *59*(2), 376–396.

Salmela, M., & von Scheve, C. (2017). Emotional roots of right-wing political populism. *Social Science Information*, *56*(4), 567–595.

Schumacher, G., Rooduijn, M., & Bakker, B. (2022). Hot populism? Affective responses to antiestablishment rhetoric. *Political Psychology*, *43*(5), 851–871.

Valentino, N. A., Hutchings, V. L., Banks, A. J., & Davis, A. K. (2008). Is a worried citizen a good citizen? Emotions, political information seeking, and learning via the internet. *Political Psychology*, *29*(2), 247–273.

van Erkel, P., & Turkenburg, E. (2020). Attitudes différentes, comportements électoraux différents, sentiments différents? [Different attitudes, different electoral behaviours, different sentiments?]. In J.-B. Pilet, P. Baudewyns, K. Deschouwer, A. Kern & J. Lefevere (Eds.), *Les Belges haussent leur voix* [The Belgians raise their voice] (pp. 153–173). Presses Universitaires de Louvain.

Vasilopoulos, P., Marcus, G. E., Valentino, N. A., & Foucault, M. (2019). Fear, anger, and voting for the far right: Evidence from the November 13, 2015 Paris terror attacks. *Political Psychology*, *40*(4), 679–704.

Vasilopoulou, S., & Wagner, M. (2017). Fear, anger and enthusiasm about the European Union: Effects of emotional reactions on public preferences towards European integration. *European Union Politics*, *18*(3), 382–405.

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, *69*, Article 102199.

Walgrave, S., Celis, K., Deschouwer, K., Marien, S., Pilet, J.-B., Rihoux, B., Van Haute, E., Van Ingelgom, V., Baudewyns, P., Van Erkel, P., Jennart, I., Lefevere, J., & Kern, A. (2022). *RepResent longitudinal survey 2019–2021*. DANS. https://doi.org/10.17026/dans-xf5-djem

Ward, D., & Tavits, M. (2019). How partisan affect shapes citizens' perception of the political world. *Electoral Studies*, *60*(4), Article 102045.

Webster, S. W. (2018). Anger and declining trust in government in the American electorate. *Political Behavior*, *40*(4), 933–964.

Webster, S. W., & Albertson, B. (2022). Emotion and politics: Noncognitive psychological biases in public opinion. *Annual Review of Political Science*, *25*(1), 401–418.

Wouters, R., De Vydt, M., & Staes, L. (2022). Truly exceptional? Participants in the Belgian 2019 Youth for Climate protest wave. *Politics of the Low Countries*. Advance online publication.

About the Authors

Luca Bettarelli is a Marie-Curie individual fellow at the Centre d'Étude de la Vie Politique (CEVIPOL, Université Libre de Bruxelles). He holds a PhD in political economics from the University of Milan-Bicocca. He has been a consultant for various international institutions, such as the World Bank and the International Monetary Fund. He is the editorial manager of the *Italian Journal of Regional Science*. His research focuses on the drivers and consequences of political polarization across individuals and territories.

Caroline Close is a professor in the Department of Political Science at the Université Libre de Bruxelles and a researcher at the Centre d'Étude de la Vie Politique (CEVIPOL). Her research interests include party politics, political behavior, representation, and democratic innovations. Her work has appeared in *Party Politics*, *Acta Politica*, *Parliamentary Affairs*, *Political Studies*, and *Representation*.

Emilie van Haute is a professor at the Université Libre de Bruxelles and a researcher at the Centre d'Étude de la Vie Politique (CEVIPOL). Her research interests focus on political parties, elections, participation, and representation. Her research has appeared in *West European Politics*, *Party Politics*, *Electoral Studies*, *Political Studies*, and *European Journal of Political Research*. She is co-editor of *Acta Politica*.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 325–335
https://doi.org/10.17645/pag.v10i4.5702

Article

Negative Party Identification and the Use of Party Cues in the Direct Democratic Context

Maxime Walder [1,*] and Oliver Strijbis [2,3]

[1] Department of Political Science and International Relations, University of Geneva, Switzerland
[2] Department of Political Science, University of Zurich, Switzerland
[3] Division of Communication, History, and Politics, Franklin University Switzerland, Switzerland

* Corresponding author (maxime.walder@unige.ch)

Submitted: 29 April 2022 | Accepted: 18 October 2022 | Published: 30 December 2022

Abstract
The use of party cues is a fundamental process of how voters adopt policy preferences. While research has shown that party identification is an important driver of political attitudes in general and policy positions in particular, we know little about how negative party identification (identifying as an opponent to a party) impacts voters' political preferences. This article aims to fill this gap in the literature by combining an experimental and observational empirical analysis of the effect of negative party identification on voters' issue preferences in the context of direct democratic decision-making. First, we analyze a survey experiment conducted during a real-world campaign on affordable housing for a popular ballot in Switzerland. Using continuous measures of party identification, we show a causal relationship between negative party identification and voters' policy preferences. Second, we use longitudinal observational data of vote choice on direct democratic policy proposals and show that voters adopt policy preferences that contrast with the policy positions of parties they oppose. In sum, the two complementary designs show that voters tend to position themselves not only in alignment with their preferred parties but also in opposition to parties with which they negatively identify. Furthermore, the results indicate that, when adopting policy preferences, negative cues may carry as much weight as positive party cues. Our analysis has important implications for understanding voters' adoption of policy preferences in general and specifically in the direct democratic context.

Keywords
direct democracy; heuristics; negative partisanship; policy position; Switzerland

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Party identification is widely regarded as one of the most important factors driving vote choice. While early scholars emphasized the idea that identification with political parties may be positive *and* negative (Campbell et al., 1960), over time, scholars devoted most of their attention to the positive aspect of party identification. Many studies highlight how this identification drives the formation of political attitudes and, consequently, the decision-making process in elections and on-ballot proposals in direct democracy (Colombo & Steenbergen, 2020). However, as scholars mainly focused their attention on the positive side of party identification, it is unclear how aversion to parties also influences voters' political attitudes.

In recent years, negative partisanship has gained prominence as a concept in electoral studies, and its

impact has been observed in the real world (see Nai et al., 2022). It has been shown, for example, that in the 2002 French presidential election, votes against the Front National candidate Jean-Marie Le Pen were massive (Medeiros & Noël, 2014). Similarly, the 2016 election famously triggered votes against Clinton for Republicans and against Trump for Democrats (Abramowitz & McCoy, 2019). In this article, we study whether negative partisanship has an independent direct effect on vote choice in direct democratic decision-making. Previous research has shown that negativity in direct democratic campaigns is widespread and has some influence on the mobilization and vote choice of individual voters (Bernhard, 2015; Nai, 2013; Nai & Sciarini, 2018). Based on this observation, we argue that individuals rely on not only positive party cues but also negative party cues, positioning themselves accordingly in opposition to parties they dislike. To our knowledge, this study is the first to look in detail at the use of negative party cues on vote choice in direct democracy.

We combine two research designs to test the existence and the use of the negative partisan heuristic in political decision-making. We consider the Swiss case and the direct democratic system where citizens and parties not only regularly take a position on complex policies, but voters also act on it. In this context, we conducted an experiment with a representative sample of 2000 Swiss citizens at the beginning of a campaign for a proposal on affordable housing. In this experiment, we asked all respondents about their support for the five main Swiss parties and subjected them to one of the parties' positions on the ballot. In a second study, we use observational data to investigate whether the result from our experimental study can be generalized across time and policy proposals. The results are affirmative and indicate that citizens do use negative party cues to position themselves on a large variety of ballot proposals, even when controlling for positive party cues. Thus, we provide evidence that voters use negative party cues to take positions on issues and that negative partisanship is an important driver of voters' attitudes.

2. Theory

Voters derive issue and policy preferences either through systematic information processing or by using heuristics (Kuklinski & Quirk, 2000; Lupia, 1994). While systematic information processing is cognitively highly demanding and requires lots of time and resources, heuristics are cognitive shortcuts that can simulate the result from a well-informed process of preference formation with low levels of information (Popkin, 1991). Since voters generally lack essential political information to form policy preferences in a complex political environment, heuristic decision-making is widespread (Achen & Bartels, 2016; Carpini & Keeter, 1996; Colombo & Steenbergen, 2020). For the adoption of issue preferences, the partisan heuristic is arguably the most prominent heuristic among the different sorts of cues voters can rely on (Kriesi, 2005). According to this heuristic, voters adopt a preference for a political issue based on their party identification. If their preferred party is in favor/against a policy proposal, voters form their attitudes in favor/against this issue position in line with the position of the party with which they identify. Previous research on party heuristics highlights their explanatory power with regard to voters' decision-making in elections and when deriving policy preferences, for instance, in the context of direct democratic votes on policy proposals (Arceneaux, 2008; Boudreau & MacKenzie, 2014; Brader & Tucker, 2012; Campbell et al., 1960; Cohen, 2003; Colombo & Kriesi, 2017; Colombo & Steenbergen, 2020; Dancey & Sheagley, 2013; Kriesi, 2005; Kuklinski & Quirk, 2000; Leeper & Slothuus, 2014; Lupia, 1994; Slothuus & Bisgaard, 2021a, 2021b).

Although there is a correlation between voters' ideology and policy preferences, studies tend to show that this correlation is driven by the cues voters take from parties. Slothuus and Bisgaard (2021b) show that party cues can temper voters' self-interest in policies, which indicates that even when voters have direct self-interest in specific policy output, they are impacted in their policy preferences by party cues. In another study, Slothuus and Bisgaard (2021a) demonstrate that voters in Denmark changed their preferences as soon as one of the main parties changed its position on a policy. This clearly shows that while voters may be able to position themselves in the ideological space, when it comes to specific policy proposals, they rely to a significant extent on the cues they receive from parties. Although this may vary between policies, party cues have at least a minimal independent effect (Slothuus & Bisgaard, 2021b). However, while studies have shown how positive party identification helps even voters with clear ideological positions derive policy preference, we do not know how negative party identification affects voters' preferences.

To a large extent, negative party identification mirrors the positive side of party identification. While positive partisanship leads to the desire for a party to win, the negative side of party identification leads to the desire for a party to lose. The source of negative party identification may arise from different world visions (Hetherington & Weiler, 2009) and, more generally, diverging ideologies (Abramowitz & Webster, 2018; Medeiros & Noël, 2014). In short, "the negative partisan might believe some people are in profound error ideologically" (Ridge, 2020, p. 5), leading them to aim for the failure of the parties that voters hold negative affect towards (Michael McGregor et al., 2015). Because voters who identify negatively with parties think these parties are ideologically wrong, they will form preferences that go against such parties' positions.

The close relation between the concept of positive and negative party identification leads to similar considerations regarding the use of positive and negative party cues. Positive party cues lead voters to adopt policy

preferences in line with the party position of the preferred parties, and negative party cues lead voters to form preferences in opposition to the disliked parties. However, the logic of the negative party heuristic can be challenged based on the existing literature on the impact of party alliances on direct democratic decision-making on ballot proposals. Indeed, Kriesi (2006) shows that the support for direct democratic proposals by a large alliance of parties increases the electoral support for said proposals. This seems to speak for a stronger impact of positive than for negative party identification since, in the latter case, we would expect that a ballot proposal would receive fewer votes when more parties are in favor of it. However, since we do not know how proposals supported or opposed by alliances of parties affect specific partisan—relative to non-partisan—voters we cannot know from this aggregate observation how strong the impact of positive vs. negative party cues is.

When two parties have the same policy position and voters support one but oppose the other, the voter will experience a clear cognitive dissonance, which reduces alignment with the preferred party—as aligning with it also means aligning with the opposing party. Thus, we expect that positive identification increases the support of the party's policy position (positive partisan cues hypothesis), and negative identification decreases the support of the party's policy position of these parties (negative partisan cues hypothesis). Figure 1 summarizes the different combinations of positions parties can have and their expected impacts on voters' policy preferences.

3. Research Designs

To test our hypotheses, we rely on two complementary studies in the context of Swiss direct democracy. Switzerland has long relied on popular ballots at every level of government: national, regional, and municipal. The direct democratic institutions of the popular initiative and referendum allow voters either to propose new constitutional features or to confirm or reject laws adopted by parliaments or municipal councils. At the national level only, Swiss voters voted on 463 ballot proposals since 1960 (Swissvotes, 2022) compared to 15 national elections, making it the most prominent form of political participation in the country. Swiss direct democracy is ideal for testing hypotheses related to the use of the partisan heuristic because such frequent popular votes provide many opportunities to study the link between parties' policy positions and voters' policy preferences. These vote choices on ballot propositions also provide a behavioral and hence particularly valid measure of policy preferences. It is not surprising, then, that the political science literature has, on several occasions, relied on the Swiss case to study partisan heuristics (Colombo & Kriesi, 2017; Kriesi, 2005).

In our study, we analyze how policy positions of the five largest Swiss parties (SP, GPS, CVP, FDP, and GPS) affect the decision-making of voters on direct democratic proposals. Although over the years, more than 20 parties have been represented, these five parties have filled between 75 and 90% of the seats in the national parliament since 1971. Focusing on these five parties ensures we cover a broad ideological spectrum, with the SP and the GPS representing the left, the FDP and SVP the liberal and conservative right, and the CVP being the center.

We combine two studies with different strengths and weaknesses to investigate the effect of negative party identification on voters' preferences regarding ballot proposals. First, we conducted an experiment with a representative sample of 2000 Swiss citizens during the

Policy positions of party with positive identification

Policy position of party with negative identification	In favor	Against
Against	**Cognitive consonnance** supporting the proposal	**Cognitive dissonance**
In favor	**Cognitive dissonance**	**Cognitive consonnance** rejecting the proposal

Figure 1. Expected effect of policy positions of parties on voters' policy preferences depending on the type of identification.

campaign of a ballot initiative on affordable housing. Second, we used historic post-vote survey data to analyze how the policy positions of parties with which the voter identifies negatively affect their preference regarding the ballot proposal. The two designs complement each other in important ways: The experiment enables us to identify the causal effect of negative party cues providing strong internal validity but is limited to a single ballot proposal. The observational design, in contrast, provides empirical evidence that this effect can be observed for different ballot proposals. Thus, while the first design provides strong internal validity but lacks ecological validity, the second aims to fill this gap by providing evidence of voters' use of negative party cues throughout the period between 1981 and 2020.

4. Experimental Evidence on Negative Party Cues

On the 9th of February 2020, the Swiss population voted on a ballot initiative on affordable housing. This initiative aimed to modify the constitution so that the state would have to intervene to build and propose more affordable housing. During the campaign, the left parties (SP and GPS) took a position in favor of the constitutional modifications proposed in the initiative, while the center and right parties (CVP, FDP, and SVP) positioned themselves against the modification. In the end, the proposal was rejected by 57.1% of the Swiss voters, with a turnout of 41.68%.

We conducted a survey experiment at the beginning of the campaign for the proposal, between the 12th of December and the 14th of January. In this survey, we interviewed a representative sample of 2,000 Swiss voters and asked them about their level of political interest, trust in government, as well as sociodemographic variables, including gender, age, and education. We then randomly assigned respondents into treatment groups in which they received information on the position of one of the five main national Swiss parties on the ballot and one control group that did not receive any additional information. The data was collected early in the campaign to ensure that the different parties did not officially position themselves on the ballot. Also, as shown in Table A6 in the Supplementary File, the treatment distribution is balanced when considering respondents' age and gender. Overall, the treatment reads as follows: "Based on the vote in the National Council, we know that *the party* is in favor/against the ballot initiative. And you, if the vote was held tomorrow, what would be your decision on the vote?" Respondents then indicated their support for the initiative with a four-point item, from *definitely yes* to *definitely no*. We added a *don't know* option and recoded the response to a binary variable indicating the respondent's support or opposition to the ballot. The control group did not see the first sentence; they just saw the second part of the question where we asked about their vote intention regarding the ballot. With the treatments, we can identify whether the information about the party position influences the voters' preferences conditionally on their identification with the party. In so doing, we follow previous research on party cues that investigated how the party position on the issue affects voter preferences (Boudreau & MacKenzie, 2014).

Following the treatment assignment and vote intention question, respondents indicated whether they see themselves (on a scale from 0 to 10) more as a *strong opponent* (0) or a *strong supporter* (10) of the five parties. In the model, we use this as a moderating variable if the treatment (position of a party) depends on the affiliation and feelings respondents have toward this party. This measure is well suited to evaluate voters' negative attitudes toward the different parties. Individuals can identify more or less and more or less positively/negatively with a group or organization. Hence, a continuous indicator is appropriate to measure identification. This said, our measure of party support allows us to distinguish between positive and negative party identification. When voters give scores under 5, they have a negative affect toward a party, and the opposite is true when they give a score above 5. A score of 5, in turn, means that a voter has neither a positive nor a negative view of the party.

To sum up, our main variables are the assignment of a party's position on a policy proposal (treatment), the support of respondents for the five main Swiss parties (moderator), and the vote intention of respondents (dependent variable). This way, we measure the conditional average treatment effect: how the effect of a party's policy position on a voter's policy preference is moderated by the support for the party. We also add the political interest of respondents and their trust in government as they directly influence the voters' decision-making, and controlling for it might therefore make the estimates more precise. Table 1 summarizes the variables we use in our model and their operationalizations.

To evaluate the treatment effect moderated by party support, we use logistic regressions and interact each treatment with the corresponding party support. This is a very restrictive modeling strategy as we consider not only the party affiliation of respondents but the specific support for five different parties. We interact the level of support of each respondent with these five parties, and we interact their support with the treatment. By considering these five interactions in the same model, we estimate how the influence of a party's policy position on voters' preference is moderated by the party identification of the party they were treated with.

It is often debated where questions used to moderate treatment effects should be placed in randomized experiments. On the one hand, measuring the moderator variables after the treatment creates the possibility of post-treatment bias (Aronow et al., 2019; Coppock, 2019; Montgomery et al., 2018). On the other hand, placing the moderator before the treatment leads to priming effects, including when these are questions on

Table 1. Description of variables used in the experiment.

Variable Name	Variable Type	Operationalization
Voter's vote intention on policy proposal	Dependent Variable	0 if *no* or *rather no*, 1 if *yes* or *rather yes*
Party's position on policy proposal	Treatment	1 if *treated*, and 0 otherwise
Party identification	Moderating Variable	0 *strong opponent*, 10 *strong supporter*
Political interest	Control variable	0 = *rather not* and *not interested*, 1 = *rather/very interested*
Trust in government	Control variable	1 = *do not trust*, 2 = *rather not trust*, 3 = *neither trust nor distrust*, 4 = *rather trust*, 5 = *completely trust*

respondents' identities (Valenzuela & Reny, 2020; Walter & Redlawsk, 2019). While research on the bias of the moderator's position in surveys is still scarce, Valentino et al. (2018) found no difference in the conditional average marginal effects with pre- and post-treatment measures. Nevertheless, scholars agree that measuring moderators within a survey experiment may lead to various causal inference issues (Sheagley & Clifford, 2022). We follow Klar et al. (2020) and Valenzuela and Reny (2020), who argue that when deciding where to place the moderator variable, there must be theoretical considerations about how biases can be minimized. With regard to this experiment, we think that priming respondents on their partisan identities would be more problematic than placing the moderator variable post-treatment, as doing so could have led to an overestimation of the treatment effect. If respondents indicate strong affection/resentment towards parties, they might be encouraged to follow/defect from the party's issue position to avoid inconsistency. In contrast, the treatment of a party's policy position is less likely to influence the measure of the moderator because inconsistency is less direct. Indeed, when respondents receive the policy position of a party, it should not substantially affect their support for it. Even though respondents align their position with that of the party they were treated with, they can still indicate an aversion for the party and be consistent as there may be various reasons to share policy positions with parties. This is less the case when considering the priming effect. Indeed, when indicating strong support for a party, it is a direct inconsistency to indicate a position that goes against the party's position. Hence, following Walter and Redlawsk (2019), we measure the moderating variable after the treatment to avoid priming respondents with their party identification.

Additional analyses also suggest that our decision to measure the moderating variable after the treatment was appropriate. As a test of whether the post-treatment bias was severe, we estimated the average marginal treatment effect of the party position on affordable housing (treatment) on party support (see Table A7 of the Supplementary File). The results show no significant treatment effect on our moderator except for the model with support for the SVP. However, even in this case, both the treatment with the SP and the SVP position shows a negative effect. This is despite the fact that these parties are clearly positioned on opposite sides of the ideological space (see Section 5, on observational evidence). While we acknowledge that this is not a definitive test, as it is impossible to test for the null hypothesis (Montgomery et al., 2018), the fact that we find no consistent relation between the treatment assignment and the moderators is reassuring.

The intuition of the model is that the party position (treatment) should have different effects on respondents who support the party whose position is seen relative to those who oppose it. In the context of the affordable housing initiative, the CVP, FDP, and SVP gave recommendations to reject the proposal, i.e., adopted a negative position. The counterfactual is no or a neutral position of the party since the respondents in the control group were not treated with the vote recommendation of the respective party, and the parties had not decided on their vote recommendation yet. Thus, we should see that supporters of the CVP, FDP, and SVP should indicate greater opposition to the ballot when they receive the treatment (positive cues hypothesis). In contrast, respondents who are opponents of these parties should adopt a vote intention more favorable to the affordable housing initiative when they receive the treatment (negative cues hypothesis).

Figure 2 presents the results of the interaction between party support and the treatment for the parties, where we stated that they were against the proposal. The results of the interactions go in the expected direction. Indeed, we observe that party supporters are more opposed to the proposal when they receive the treatment, and party opponents are more supportive of the proposal. Interestingly, the direction of the relationship between support for the SVP and voters' preferences on the proposal changes for treated and untreated respondents. While there is a positive relationship between

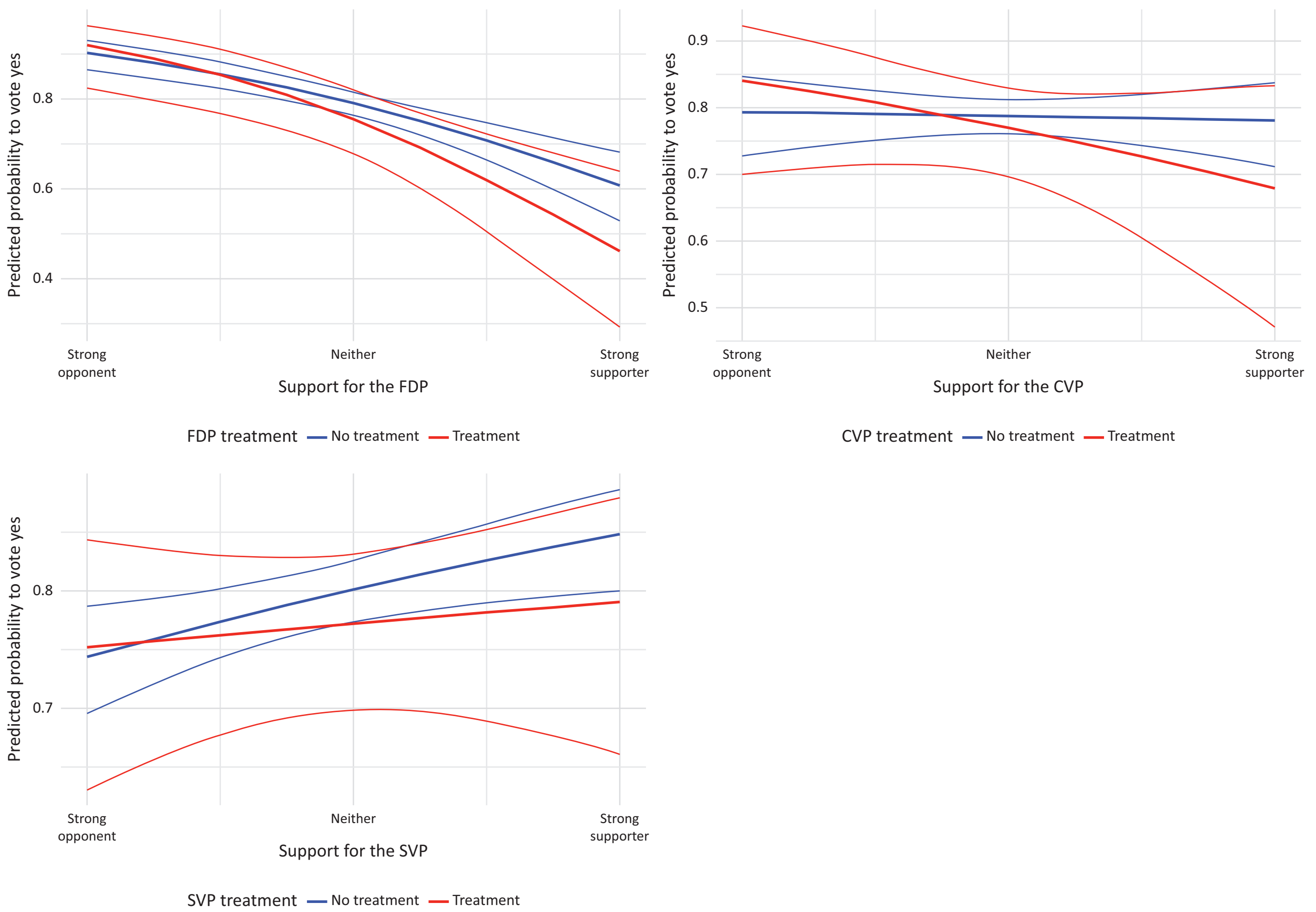

Figure 2. Moderating effect of party support and the treatment on voters' preferences on the affordable housing initiative for parties who are against the ballot proposal (FDP, CVP, and SVP).

support for the SVP and support for the ballot proposal in the control condition, this relationship turns negative for respondents who received the treatment. However, for every party, the size of the confidence interval also suggests that this effect is rather small and not significant. Thus, while Figure 2 shows that the interactions between parties' support and the treatment go in the expected direction, we do not find significant effects for parties who positioned against the proposal—a point to which we will return below.

We now turn to the treatment of parties who positioned in favor of the affordable housing initiative (SP and GPS). Figure 3 presents the interaction effects between the treatment of parties in favor of the initiative

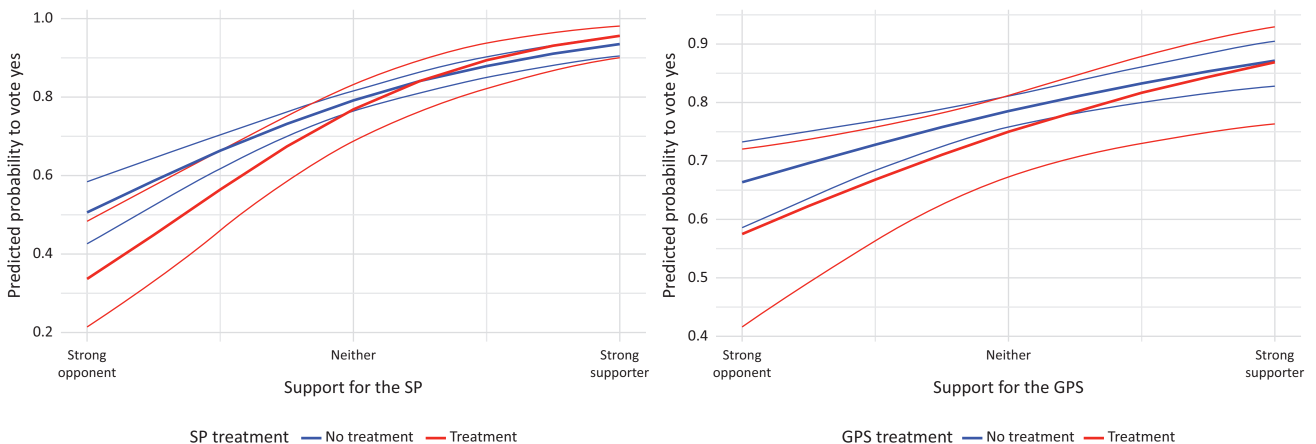

Figure 3. Moderating effect of party support and the treatment on voters' preferences on the affordable housing initiative for parties who are in favor of the ballot proposal (SP and GPS).

and the support for the party on voters' preference for the proposal. It shows that these interactions also go in the expected direction. Indeed, opponents are more against the proposal, and support is higher among the party supporters who received the treatment. The effect also appears substantial if we compare the effect size of positive and negative party cues, where we see that voters who oppose the SP or the GPS were more impacted by the vote recommendation than voters who support them. Despite the strong relationship between party identification and voters' preference, we find evidence that treating respondents with parties' position influences the formation of voters' preferences, conditional to their identification with the party.

Overall, the results of the experiment show that the causal effect of parties' positions on voters is not uniform and is moderated by the level of support for the party. This provides evidence that voters use both positive and negative party cues to take a position on policies.

Although all interaction terms go in the expected direction, different factors may explain the lack of significance of the conditional treatment effects. First, our modeling strategy is very restrictive, and party support is already a strong predictor of voters' positions. Second, the ballot proposal for affordable housing can be easily linked to ideology as it directly relates to the economic inequality issue. Hence, although parties had not yet communicated their official positions on the ballot, voters may have been able to guess them. This said, although most of our results are not significant at conventional levels, all conditional average marginal treatment effects go in the expected direction. To evaluate whether this result can be generalized, we turn now to the study with observational data.

5. Evidence on Negative Party Cues From Observational Data

To test our hypotheses on many policy proposals, we merged two datasets. First, the VoxIt data (Kriesi et al., 2017) contains post-survey data on 297 ballot proposals that the Swiss population voted on between 1981 and 2016. Second, the VOTO (FORS, 2020) data consists of post-vote surveys of 13 ballot proposals that took place between 2017 and 2020. Thus, we analyze the effect of party vote recommendations on the vote choice of party voters over more than 300 ballot proposals.

The data contains information on party vote recommendation (our measure for parties' policy positions), respondents' party affiliation, and vote choice, among others. However, in contrast to the experiment, we do not know how respondents feel toward parties with which they do not identify. To operationalize negative party identification, we use results from the experiment and measure the correlation between the support for the different parties. Figure A1 in the Supplementary File presents the correlation between the different measures of party support and shows that the correlation between support for the different parties follows a left–right divide. Indeed, we see that there are strong negative correlations between the GPS and the SP (on the left) and the SVP and FDP (on the right). As a center-right party, support for the CVP is not strongly negatively correlated with support for any other party. These divisions between Swiss parties have deep historical roots. Glass (1978) already provided evidence that the SP, the FDP, and the SVP positioned themselves on opposite sides of the ideological space in 1972. More recent work of Hug and Schulz (2007) has shown that the ideological positions of Swiss parties are very stable over time. Scholars have also emphasized that—despite a certain shift in the cleavage structure (Kriesi, 2015)—the main division among Swiss parties has remained stable since 1960 and that the left has concentrated around the SP and GPS (Durrer de la Sota et al., 2021). Thus, voters who identify with one of the two main left parties are likely to negatively portray the two most prominent parties of the right (the FDP and SVP).

To provide empirical evidence on the long-lasting division between left and right Swiss parties, we use the historical survey data of the observable study (FORS, 2020; Kriesi et al., 2017). First, Figure A3 of the Supplementary File plots the average left–right self-position of voters that identify with the different parties over time. It shows that the divide between Swiss voters who identify with the left and right parties is constant. As identification is strongly related to ideology, we think it is reasonable to posit that voters who positively identify with one of the left/right parties are likely to identify negatively with a party on the other ideological side, not only during the affordable housing initiative—for which we could observe it—but more generally. This is also what the parties' positions on direct democratic proposals suggests. Indeed, Figure A2 in the Supplementary File shows the share of ballot proposals on which the four different parties took the same policy position. The figure shows that while the SP and the GPS share the same policy position on ballots more than 90% of the time, they often do not share their position with the right parties. This shows that voters who identify with the left and right parties not only position themselves at the other end of the left–right scale, but that they also experience a strong division between these two blocks with regard to their policy positions in direct democratic votes. We thus derive that voters who identify with the SP and GPS are more likely to see themselves as opponents of the FDP and SVP and voters who identify with the FDP and SVP are likely to see themselves more as opponents of the SP and the GPS. Negative party identification is largely driven by ideological divergence. As we show, the ideological divergences between left and right Swiss parties have deep historical roots. We thus analyze how right/left voters form preferences using left and right parties' positions. Although this operationalization of negative party identification has clear limitations, we think that these are compensated for by the important advantages of this very large dataset.

We estimate how voters' preferences who identify with left or right parties are influenced by the position of the left and right parties on the issue. For instance, we test how the position of SP and SVP voters are influenced by the interaction of the SVP and SP positions as well as the party the respondents' support. We only test the moderating effect of pairs of parties' policy positions. For instance, a model that estimates the moderating effect of the SP and SVP positions limits the analyses to SP and SVP voters. Finally, we control for the policy positions of the five main parties (CVP, FDP, SVP, SP, and GPS) independent of the interaction terms, and the strength of parties based on the share of seats in the national parliament. We control for the latter as it influences the potential threat they represent and the institutional type of ballot (initiative, facultative referendum, and mandatory referendum), as these institutions affect the level of support for ballot proposals.

The model evaluates the effect of opposing parties' positions on proposals relative to each other. We thus run four models for all the combinations identified: SP versus SVP, SP versus FDP, GPS versus SVP, and GPS versus FDP. We first show the two models that include the SP in the interaction, followed by the models with the GPS included in the interaction.

Figure 4 presents the results of the interaction between the positions of the SP and the SVP (on the left) and the SP and the FDP on the right. We see that the SP's position on ballots has no effect on the preferences of SVP voters when the SVP takes a position against the proposal. However, when the SVP positions in favor of the ballot, their voters' preferences are moderated by the SP's position, i.e., they are more opposed to the proposal when the SP supports it.

Similarly, we see that SP voters oppose more proposals that are supported by the SVP when the SP positions against them. If we look at the right side of the figure, we see that while the FDP position does not affect the preferences of SP voters, the opposite is not true. Indeed, FDP voters align less with the FDP position when the SP shares the same position than when the FDP and SP have diverging positions.

Although the moderating effect of the opposing party's position on voters' policy preferences is not uniform, Figure 4 shows that SP voters oppose SVP recommendations and SVP voters adopt preferences against those of the SP. Also, it shows that while SP voters are not affected by FDP positions on ballot proposals, FDP voters oppose significantly more proposals when the SP supports them than when it positions itself against them.

Figure 5 plots the moderating effect of the GPS's positions, and the SVP or FDP positions. First, on the left side of the figure, we see that SVP voters react negatively to GPS positions. SVP voters oppose significantly more proposals when the GPS positions in favor of them. For GPS voters, we see that the SVP position on ballot proposals does not affect their preferences when the GPS supports the proposal. However, they react negatively to the SVP's position when the GPS opposes the proposal. Indeed, in this case, they are even more against the proposal than when the SVP is in favor of it.

On the right side of the figure, we see that the FDP and GPS voters react negatively to the other party's position. Indeed, GPS voters oppose proposals more when the FDP is in favor of them than when the FDP opposes them. Also, FDP voters oppose the proposal more when the GPS supports it than when it opposes it.

Overall, the evidence based on observational data presented in the second study of this article supports our hypotheses and shows that voters adopt preferences in opposition to the position of opposing parties. Left-party voters tend to take a position against the right parties and *vice versa*.

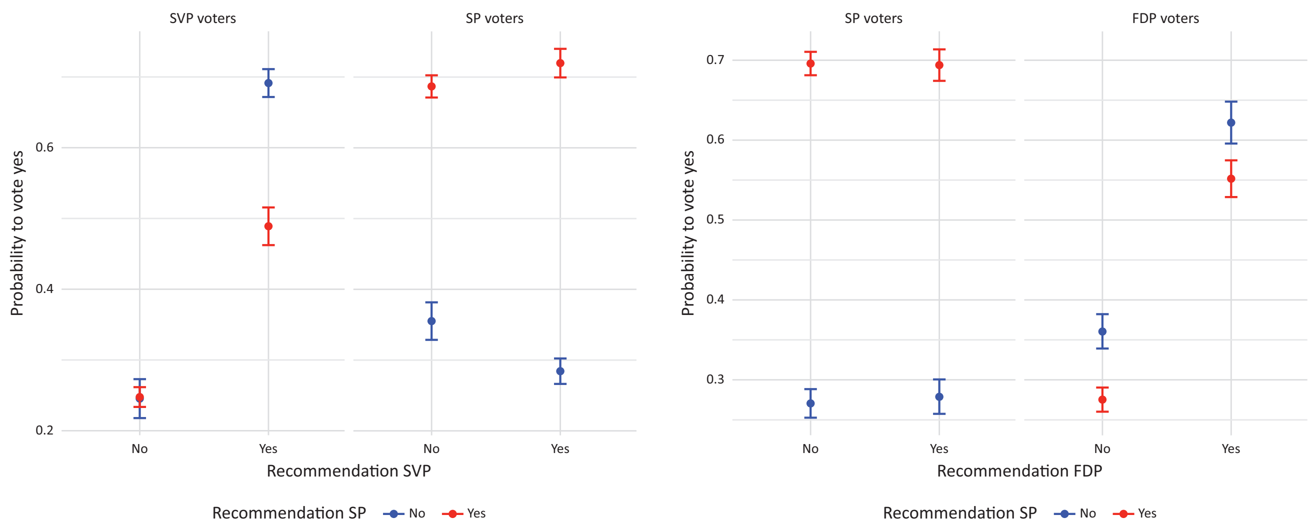

Figure 4. Moderating effect of the policy position of the SP and SVP (left) and FDP (right) on voters' positions on ballot proposal by party affiliation.

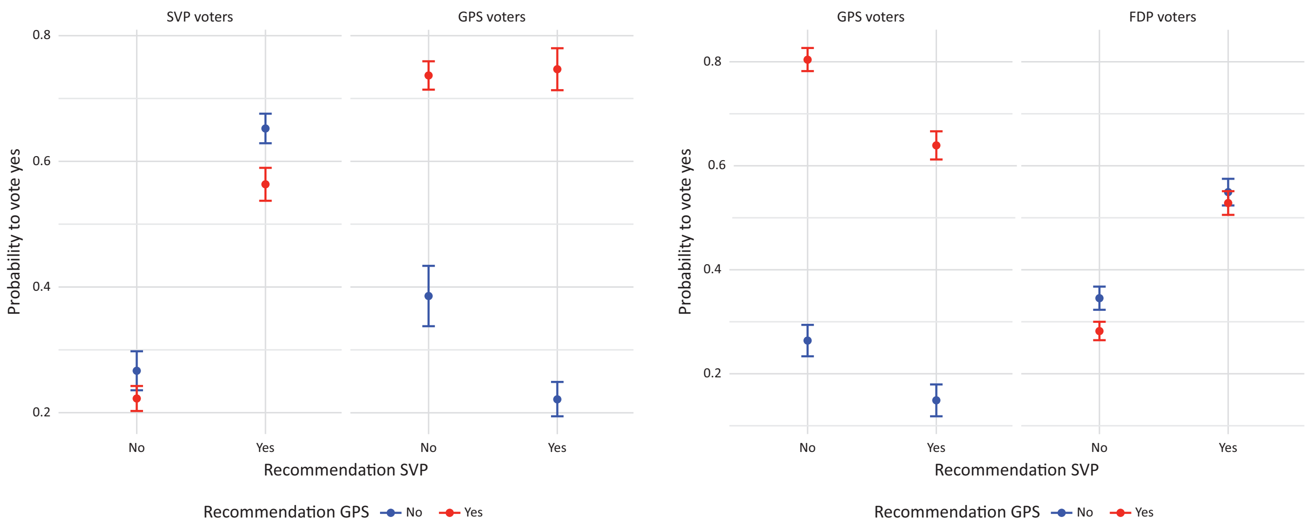

Figure 5. Moderating effect of the policy position of the GPS and SVP (left) and FDP (right) on voters' preferences on ballot proposals by party affiliation.

6. Conclusion

This article studies the use of negative party cues on voters' policy preferences. Using experimental and observational evidence, we show that the use of negative cues drives the decision-making process in the direct democratic setting. We also show that it has a complementary explanatory power to the use of positive party cues. These results have consequences for the role that negative partisanship plays in the decision-making process.

However, the consequence of negative party identification may be even larger than has been discussed so far in this article. Indeed, our results suggest that core party voters will align less with the position of the party they identify with in the event of a large alliance of parties for a ballot proposal. We show that the support for the left and right party positions by voters who identify with these parties is higher when the parties have opposite positions. Thus, parties with opposing views may have an electoral disadvantage in defending a common position. Indeed, in this case, we show that the support of the party position by the core voters is lower than when opposing parties' positions are in opposition to each other. The use of negative party cues—and negative partisanship in general—may have a detrimental effect on the formation of party coalitions in democracies as it may give dissonant cognitive information to parties' core electorate and generate vote defection in subsequent elections.

Our article presents evidence that supports the fundamental aspects of negative partisanship in voters' attitude formation. However, several aspects limit our ability to draw definitive conclusions on the importance of negative party cues. First, our experimental design is limited to a single policy proposal with a clear left–right divide. We think that future studies should conduct experiments on ballots with different ideological divisions to deepen our understanding of the importance of negative party cues. Second, our observational study relies on a crude operationalization of negative partisanship because, in our context, there were no surveys available that measured negative party identification. Hence, in order to gain additional insight into the role of negative party identification, political surveys should systematically include questions that enable research to have precise operationalization of the negative side of partisanship. This would complement recent studies—with this special issue as a prominent example—that show that the negative side of party identification is an essential component of various aspects of party competition, voting behavior, and the quality of democracy in general.

Acknowledgments

We would like to thank the editors of this thematic issue and three anonymous reviewers for their very helpful comments.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Abramowitz, A., & McCoy, J. (2019). United States: Racial resentment, negative partisanship, and polarization in Trump's America. *The ANNALS of the American Academy of Political and Social Science*, *681*(1), 137–156.

Abramowitz, A. I., & Webster, S. W. (2018). Negative partisanship: Why Americans dislike parties but behave like rabid partisans. *Political Psychology*, *39*, 119–135.

Achen, C. H., & Bartels, L. (2016). *Democracy for realists: Why elections do not produce responsive government*. Princeton University Press.

Arceneaux, K. (2008). Can partisan cues diminish democratic accountability? *Political Behavior*, *30*(2), 139–160.

Aronow, P. M., Baron, J., & Pinson, L. (2019). A note on dropping experimental subjects who fail a manipulation check. *Political Analysis*, *27*(4), 572–589.

Bernhard, L. (2015). Going negative in direct-democratic campaigns. In A. Nai & A. Walter (Eds.), *New perspectives on negative campaigning* (pp 145–63). ECPR Press.

Boudreau, C., & MacKenzie, S. A. (2014). Informing the electorate? How party cues and policy information affect public opinion about initiatives. *American Journal of Political Science*, *58*(1), 48–62.

Brader, T., & Tucker, J. A. (2012). Following the party's lead: Party cues, policy opinion, and the power of partisanship in three multiparty systems. *Comparative Politics*, *44*(4), 403–403.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. University of Chicago Press.

Carpini, M. X. D., & Keeter, S. (1996). *What Americans know about politics and why it matters*. Yale University Press.

Cohen, G. L. (2003). Party over policy: The dominating impact of group influence on political beliefs. *Journal of Personality and Social Psychology*, *85*(5), 808–822.

Colombo, C., & Kriesi, H. (2017). Party, policy—Or both? Partisan-biased processing of policy arguments in direct democracy. *Journal of Elections, Public Opinion and Parties*, *27*(3), 235–253.

Colombo, C., & Steenbergen, M. R. (2020). Heuristics and biases in political decision making. In P. Redlawsk (Ed.), *Oxford encyclopedia of political decision making*. https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-974

Coppock, A. (2019). Avoiding posttreatment bias in audit experiments. *Journal of Experimental Political Science*, *6*(1), 1–4.

Dancey, L., & Sheagley, G. (2013). Heuristics behaving badly: Party cues and voter knowledge. *American Journal of Political Science*, *57*(2), 312–325.

Durrer de la Sota, C., Gethin, A., & Martínez-Toledano, C. (2021). *Party system transformation and the structure of political cleavages in Austria, Belgium, the Netherlands and Switzerland, 1967–2019* (Working Paper No. 2021/08). World Inequality Lab.

FORS. (2020). *VOTO studies: Standardized post-vote surveys, 2016–2020* [Data set]. SWISSUbase. https://doi.org/10.23662/FORS-DS-1231-1

Glass, H. E. (1978). Consensus and opposition in Switzerland: A neglected consideration. *Comparative Politics*, *10*(3), 361–372.

Hetherington, M. J., & Weiler, J. D. (2009). *Authoritarianism and polarization in American politics*. Cambridge University Press.

Hug, S., & Schulz, T. (2007). Left–right positions of political parties in Switzerland. *Party Politics*, *13*(3), 305–330.

Klar, S., Leeper, T., & Robison, J. (2020). Studying identities with experiments: Weighing the risk of posttreatment bias against priming effects. *Journal of Experimental Political Science*, *7*(1), 56–60.

Kriesi, H. (2005). *Direct democratic choice: The swiss experience*. Lexington Books

Kriesi, H. (2006). Role of the political elite in swiss direct-democratic votes. *Party Politics*, *12*(5), 599–622.

Kriesi, H. (2015). Conclusion: The political consequences of the polarization of Swiss politics. *Swiss Political Science Review*, *21*(4), 724–739.

Kriesi, H., Brunner, M., & Lorétan, F. (2017). *Enquêtes standardisées VoxIt 1981–2016* [Standardized VoxIt surveys 1981–2016] [Data set]. SWISSUbase. https://doi.org/10.23662/FORS-DS-688-2

Kuklinski, J. H., & Quirk, P. J. (2000). Reconsidering the rational public: Cognition, heuristics, and mass opinion. In A. Lupia, M. D. McCubbins, & S. L. Popkin (Eds.), *Elements of reason: Cognition, choice, and the bounds of rationality* (pp. 153–182). Cambridge University Press.

Leeper, T. J., & Slothuus, R. (2014). Political parties, motivated reasoning, and public opinion formation. *Political Psychology*, *35*, 129–156.

Lupia, A. (1994). Shortcuts versus encyclopedias: Information and voting behavior in California insurance reform elections. *American Political Science Review*, *88*(1), 63–76.

Medeiros, M., & Noël, A. (2014). The forgotten side of partisanship: Negative party identification in four Anglo-American democracies. *Comparative Political Studies*, *47*(7), 1022–1046.

Michael McGregor, R., Caruana, N. J., & Stephenson, L. B. (2015). Negative partisanship in a multi-party system: The case of Canada. *Journal of Elections, Public Opinion and Parties*, *25*(3), 300–316.

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How conditioning on posttreatment variables can ruin your experiment and what to do about it. *American Journal of Political Science*, *62*(3), 760–775.

Nai, A. (2013). What really matters is which camp goes dirty: Differential effects of negative campaigning on turnout during Swiss federal ballots. *European Journal of Political Research*, *52*(1), 44–70.

Nai, A., Garzia, D., Aaldering, L., da Silva, F. F., & Gattermann, K. (2022). For a research agenda on negative politics. *Politics and Governance*, *10*(4), 243–246.

Nai, A., & Sciarini, P. (2018). Why "going negative?" Strategic and situational determinants of personal

attacks in Swiss direct democratic votes. *Journal of Political Marketing*, *17*(4), 382–417.

Popkin, S. (1991). *The reasoning voter*. University of Chicago Press.

Ridge, H. M. (2020). Enemy mine: Negative partisanship and satisfaction with democracy. *Political Behavior*, *44*, 1271–1295. https://doi.org/10.1007/s11109-020-09658-7

Sheagley, G., & Clifford, S. (2022). *No need for a panel study: Measuring moderators doesn't alter treatment effects*. Unpublished Manuscript.

Slothuus, R., & Bisgaard, M. (2021a). How political parties shape public opinion in the real world. *American Journal of Political Science*, *65*(4), 896–911.

Slothuus, R., & Bisgaard, M. (2021b). Party over pocketbook? How party cues influence opinion when citizens have a stake in policy. *American Political Science Review*, *115*(3), 1090–1096.

Swissvotes. (2022). *Swissvotes—Die datenbank der eidgen ̈ossischen volksabstimmungen* [Swiss votes—The database of federal referendums] [Data set]. www.swissvotes.ch

Valentino, N. A., Neuner, F. G., & Vandenbroek, L. M. (2018). The changing norms of racial political rhetoric and the end of racial priming. *The Journal of Politics*, *80*(3), 757–771.

Valenzuela, A. A., & Reny, T. T. (2020). Evolution of experiments on racial priming. In J. N. Druckman & D. P. Green (Eds.), *Advances in experimental political science* (pp. 447–467). Cambridge University Press.

Walter, A. S., & Redlawsk, D. P. (2019). Voters' partisan responses to politicians' immoral behavior. *Political Psychology*, *40*(5), 1075–1097.

About the Authors

Maxime Walder is a post-doctoral researcher at the University of Geneva, Switzerland. His doctoral dissertation focused on the link between parties and voters. He now works with a National Research Project in Switzerland on the effect of digital transformation on the formation of political attitudes. His main research interests are related to the study of political behavior, the public opinion of voters and political elites, and the formation of political attitudes.

Oliver Strijbis is a Swiss National Science Foundation assistant professor at the University of Zurich and an affiliated professor for political science at Franklin University Switzerland. He is the leader of the research project The Effect of Campaign Events on Direct Democratic Choices: Evidence From Prediction Markets. His main fields of interest are comparative politics, political behavior, ethnicity, migration, and prediction markets. He is an editor of *The Struggle Over Borders: Cosmopolitanism and Communitarianism* (Cambridge University Press) and has published in the *European Journal of Political Research*, *European Political Science Review*, *Electoral Studies*, *Party Politics*, *Sociological Methods and Research*, *International Journal of Forecasting*, *Ethnic and Racial Studies*, and *Journal of Ethnic and Migration Studies*, among others.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 336–348
https://doi.org/10.17645/pag.v10i4.5756

Article

Fueling Toxicity? Studying Deceitful Opinion Leaders and Behavioral Changes of Their Followers

Puck Guldemond [1,*], Andreu Casas Salleras [2], and Mariken van der Velden [2]

[1] Strategic Communication Group, Wageningen University & Research, The Netherlands
[2] Department of Communication Science, Free University Amsterdam, The Netherlands

* Corresponding author (puck.guldemond@wur.nl)

Submitted: 3 May 2022 | Accepted: 26 August 2022 | Published: 30 December 2022

Abstract
The spread of deceiving content on social media platforms is a growing concern amongst scholars, policymakers, and the public at large. We examine the extent to which influential users (i.e., "deceitful opinion leaders") on Twitter engage in the spread of different types of deceiving content, thereby overcoming the compartmentalized state of the field. We introduce a theoretical concept and approach that puts these deceitful opinion leaders at the center, instead of the content they spread. Moreover, our study contributes to the understanding of the effects that these deceiving messages have on other Twitter users. For 5,574 users and 731,371 unique messages, we apply computational methods to study changes in messaging behavior after they started following a set of eight Dutch deceitful opinion leaders on Twitter during the Dutch 2021 election campaign. The results show that users apply more uncivil language, become more affectively polarized, and talk more about politics after following a deceitful opinion leader. Our results thereby underline that this small group of deceitful opinion leaders change the norms of conversation on these platforms. Hence, this accentuates the need for future research to study the literary concept of deceitful opinion leaders.

Keywords
computational communication science; disinformation; opinion leaders; social media; the Netherlands; Twitter

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Social media (SM) platforms play a key role in our daily lives. People increasingly use SM to interact with friends and family, voice their opinions, consume news, and engage in politics (Popan et al., 2019; Spohr, 2017; Weeks et al., 2017). However, some information on SM is misleading, i.e., untrue, partly true, and potentially purposefully deceitful. This has been studied by scholars using the concepts of fake news (Egelhofer & Lecheler, 2019; Guess et al., 2019), disinformation (Bennett & Livingston, 2018; McKay & Tenove, 2021), rumors (DiFonzo & Bordia, 2007; Friggeri et al., 2014), or conspiracies (Douglas et al., 2019; Sunstein & Vermeule, 2009), amongst other concepts. Academics have raised concerns, stating that deceitful content endangers democracy and society at large (Groshek & Koc-Michalska, 2017; Lazer et al., 2018; McKay & Tenove, 2021). For example, deceitful content has led to online discussions between SM users in which uncivil language is common, especially when these discussions are about polarizing political topics (Weeks & Gil de Zúñiga, 2021). This tone of voice, in turn, fuels toxicity on SM platforms (Kim et al., 2021). That is, uncivil language spills over to other SM users, which affects their attitudes towards those who are addressed in these messages, potentially leading to polarization.

Uncivil messages are those that contain curse words, are insulting, harassing, very dismissive towards others, racist or against a minority group, or are misogynistic, enabling a toxic sphere (Davidson et al., 2017; Theocharis et al., 2016). To remedy toxicity on their platforms, keeping them a healthy place for public debate, SM companies often remove users who spread deceitful content. This removal has fueled a societal debate about whether these actions are justified because such regulatory measures stand in contrast to the claims that SM would provide more equal opportunities for the free expression of political views than traditional media (Balkin, 2017). Hence, SM companies as private actors are engaged in regulating the "practical conditions of speech" in the digital space (Balkin, 2017). Yet, are those who spread deceitful content harmful to others? Currently, the empirical evidence on if and to what extent deceitful content harms other platform users is scarce. Therefore, we need a systematic study on disseminating a variety of types of deceitful content (e.g., fake news, conspiracies, rumors, and disinformation) and the effect thereof on other SM users.

We argue that the current state of the field aiming to understand the negative consequences of deceitful content is limited in three ways. First, previous research has been very compartmentalized. Scholars have studied different types of deceitful content in isolation (Weeks & Gil de Zúñiga, 2021). We, however, claim that when disseminating deceitful information, SM accounts spreading deceitful content often do not stick to just one type of deceitful information: They spread a variety of deceitful content throughout. Anecdotally, the now purged SM accounts of far-right radio show host Alex Jones show that he engages in conspiracies, as well as rumors and misleading information (Berr, 2019; Coaston, 2018; Haselton, 2019; Paul, 2019; Rosdorff, 2018). In our empirical analysis, we assess the validity of our claim that these kinds of salient accounts engage in the spread of different types of deceiving information. Thereby, this study meets and expands the work of Weeks and Gil de Zúñiga (2021), who call for research that goes beyond the mere distinction between different types of deceitful information. Furthermore, we build upon the work of Chadwick and Stanyer (2022), who theoretically argue for the need to have an overarching framework bridging the myriad of studies addressing deceitful content. We theorize and empirically demonstrate how various types of deceitful content are addressed, allowing us to gauge the harm of this content to other users and, thereby, to democracy and society at large. Second, existing research is focused on the type of content spread rather than on the SM accounts disseminating this information. If we aim to better understand the effect these salient accounts have on other users and, to some extent, whether the removal of accounts disseminating deceitful content is justified, we argue that not the content but the SM accounts should be at the center of analysis. We are interested in the negative effects that all these types of deceitful content have on SM users, not just a particular type of deceitful content. Building upon the two-step flow of communication theory (Katz & Lazarsfeld, 1955), we introduce a theoretical concept that examines these salient accounts fueling SM platforms with toxicity by posting deceitful content, terming them "deceitful opinion leaders" (DOLs). Third, while there is plenty of existing knowledge about the overall prevalence and dissemination of deceitful content, we know little about the effects that DOLs have on their followers and other users on SM platforms. Scholars suggest that exposure to deceitful messages can have harmful consequences, such as adopting more uncivil behavior, and lead to increasing levels of affective polarization (Iyengar et al., 2012, 2019; Popan et al., 2019; Theocharis et al., 2016; Yarchi et al., 2021). We test whether exposure to deceitful information actually has these malicious effects.

To empirically assess the type of deceitful content DOLs spread as well as its effect on other SM users, we use an innovative research design that allows us to study the type of content DOLs disseminate and if this fuels toxicity on Twitter. Twitter is known as a key mainstream platform that allows us to collect the data needed to test our hypotheses. For a two-month period (March 2, 2021, till May 4, 2021), we tracked eight Dutch DOLs (Maurice de Hond, Lange Frans, Sietske Bergsma, Robert Jensen, Blackbox News, Wierd Duk, Cafe Weltschmerz, and Isa Kriens) and their followers. These DOLs are not an exhaustive nor representative list of DOLs in the Dutch Twittersphere. Yet, they are well known for engaging in the dissemination of deceitful information (e.g., see "YouTube verdedigt verwijderen account Lange Frans," 2020), and thereby a most likely case to test our approach and theoretical concept. All DOLs have accounts with a high number of followers (i.e., more than 11,000), showing that these DOLs voice opinions that are valued and accepted by others. Moreover, DOLs often spread deceitful content about highly polarized and political issues. This results in an (online) public space fueled with toxicity and deceitful content (Bergmann, 2020). The collected messages of these DOLs allow us to assess the validity of our claim that DOLs engage in the spread of different types of deceiving content. For each day in the period under investigation, we monitored each DOL for if they had new followers (N = 32,245). Subsequently, for each of these new followers, we collect the tweets they posted before and after they started following a DOL. Our analysis is two-fold. First, we look at the tweets posted by the DOLs and use content analysis to corroborate that they indeed engage in a wide variety of deceitful content, such as rumors and disinformation. Then, we look at the tweets sent by the new followers before and after and use computational methods to test the extent to which they become more politically engaged and post more uncivil and affectively polarized messages after following a DOL.

We show that, after starting to follow one of the eight DOLs in our sample, these users did increase their

number of political, uncivil, and affectively polarized tweets. The effects are statistically significant and of substantial magnitude. We observe stronger longer-term (30 days) than shorter-term (15 days) effects, although, after two weeks, their behavior starts reverting to levels similar to those before following the DOL. We also observe stronger effects for those who started following more than one DOL. Our results thereby underline that while there is a small group of DOLs, they do have a substantial effect on how other SM users behave on these SM platforms. To keep SM platforms a healthy forum for public debate, SM companies regulate what can be posted. Fueled by fear that the dissemination of deceitful information distorts a healthy public debate and, thereby, is detrimental to society, SM accounts engaging in this behavior are purged. Our results, however, demonstrate that following a DOL has a gateway effect: Not only are SM users adopting their norms of conversation (i.e., using more uncivil language), but they also introduce their SM followers to a view of politics that these followers feel more comfortable to engage in. This sheds important light on the question of how to regulate SM platforms so that they can maintain fostering public debate without endangering the democratic process of deliberation.

2. Deceitful Opinion Leaders on Social Media and Their Effects

Over the last decades, the media environment has changed drastically into a high-choice media environment (Van Aelst et al., 2017). This has affected the communication flow from the media to the masses. Many people receive news via SM through one of their online connections (Weeks et al., 2017). Hence, these connections function as a mediator between the media and the mass public. This process was first explained by Katz and Lazarsfeld (1955) as the two-step flow of communication theory, which acknowledges this process of person-to-person influence and calls these mediators opinion leaders. Those are people that are held in high esteem and whose opinions are valued and accepted by others (Bergström & Jervelycke Belfrage, 2018; Choi, 2015; Katz & Lazarsfeld, 1955). In the early days of mass media, opinion leaders received information from the media and shared that information with their network via (offline) personal interactions. In the digital age, this process is similar but takes place in an online environment: SM users seek out certain individual SM accounts for guidance and information (Choi, 2015). The information that SM users are exposed to depends on the opinions, interests, and behavior of their online connections (Bergström & Jervelycke Belfrage, 2018). These opinion leaders inform and thereby potentially shape the attitudes of less active recipients (Bergström & Jervelycke Belfrage, 2018; Carlson, 2019). Yet, they do not necessarily need to be message carriers for the greater good. In recent years, we have witnessed opinion leaders that deliberately spread information that is untrue or deceiving, such as Alex Jones or Lange Frans in the Dutch context. Influential accounts that engage in this behavior we coin as DOLs. DOLs are defined as SM users (a) with a large number of followers and (b) who engage in the production and dissemination of at least one type of deceitful content to their audiences.

Why do people follow DOLs, and what is the effect thereof? Previous research demonstrates that most people are not necessarily engaged with politics, but do enjoy following entertaining content. As a by-product of seeking entertainment, politically inattentive individuals are exposed to information about political and societal issues (Baum, 2002). Social networks like Twitter provide increasing opportunities for people to be exposed to political content, even when using Twitter for different purposes, such as entertainment (Kim et al., 2013). DOLs typically post highly entertaining and engaging content, such as sarcastic or cynical comments. Hence, people are, in part, likely to follow them for entertainment value. A side-effect of following DOLs is that their followers are *incidentally exposed* (Bergström & Jervelycke Belfrage, 2018; Kim et al., 2013; Weeks et al., 2017) to political content—i.e., when DOLs tweet about societal, controversial, and political issues, their followers (and the followers of their followers via sharing patterns) see this content. The same dynamic holds for exposure to misleading information (Lazer et al., 2018; Stroud, 2008). We argue that DOLs have a key role in the information others receive, resulting in a high influence on what DOL followers talk about (Zaller, 1992). That is, the topics of conversation—i.e., the deceitful information about societal and political topics—likely spillover to the DOL followers, leading to the following hypothesis:

> H1: After following DOLs, users will tweet more about politics than they did before following them.

Next to *what* DOLs talk about, *how* they speak about political topics is also likely to be carried over to their followers. According to Weeks and Gil de Zúñiga (2021), online political interactions are often uncivil. The highly emotional nature of SM platforms provides a "perfect storm" for the spread of deceiving and misleading content (Weeks & Gil de Zúñiga, 2021). DOLs often use inflammatory and uncivil rhetoric when discussing political topics or when referring to politicians (for an example, see Table 3). Due to anonymity, the threshold for uncivil behavior is lowered on SM platforms (Groshek & Koc-Michalska, 2017; Theocharis et al., 2016). In an SM environment, people tend to say and do things that they would not necessarily do when being in the offline world (Suler, 2004). Therefore, SM platforms facilitate this uncivil behavior online (Groshek & Koc-Michalska, 2017). This, in turn, results in the usage of more uncivil language, posing threats, hard criticism, and showing anger and hatred on SM platforms online, creating an online sphere rife with uncivil behavior (Suler, 2004;

Theocharis et al., 2016). Impolite and uncivil discourse on SM platforms has a poisonous and polarizing effect. When people are exposed to incivility, they are more likely to use incivility in their comments and messages (Gervais, 2015; Theocharis et al., 2016). This implies that those following DOLs, who are expected to use uncivil and inflammatory language, are more likely to mimic their rhetorical style, leading to the following hypothesis:

> H2: After following DOLs, users will utilize more uncivil language.

Uncivil behavior on SM platforms reduces openness towards outgroups, as uncivil discourse has poisonous and polarizing effects (Groshek & Koc-Michalska, 2017; Theocharis et al., 2016). As mentioned above, DOLs often talk about political topics or politicians in an uncivil manner. They use an "us versus them" rhetoric when referring to the political elite. By doing so, they create an in-group (DOLs and their followers) and an out-group (the political elite and their followers). Based on the social identity theory (Tajfel & Turner, 1979), scholars have theorized and demonstrated that belonging to an in-group with a strong social identity leads to the disliking and disfavoring of out-groups (Harteveld, 2021; Iyengar et al., 2012, 2019). Online, this results in SM users following more like-minded accounts that fit within their in-group. This implies that once SM users follow a DOL, they are likely to be immersed in an online community of like-minded people, forming online homogenous networks (Barberá, 2015; Barberá et al., 2015; Shu et al., 2017). These homogeneous social networks reduce the tolerance for alternative worldviews and amplify affective polarization, resulting in division and animosity between different parties, individuals, or groups that hold opposite views on (political) topics (Iyengar et al., 2012, 2019; Lazer et al., 2018; Yarchi et al., 2021). Assuming that SM users following and mimicking a DOL can be seen as a united front (i.e., in-group), they are likely to view others as an out-group whom they oppose. Thereby, they presumably contribute to rising hostility toward other societal groups. By following DOLs, we expect users to become less tolerant, hence more polarized, towards outgroups with different opinions and ideas. Therefore, we expect the following:

> H3: After following DOLs, users will become more affectively polarized.

3. Data and Methods

We collected the following data to assess the extent to which DOLs engage in the dissemination of different kinds of deceitful content, as well as to test our three hypotheses about the effects that they have on the behavior of their followers. First, we selected a set of DOLs to study. Then, we tracked their SM behavior to explore the types of deceitful content they posted. In addition, we needed to track the SM behavior of their followers. Ideally, for clear identification, we wanted to track and study their behavior before versus after they started following a given DOL.

For a two-month period during the 2021 Dutch elections (March 2 through May 4, 2021), we studied the Twitter behavior of a convenience sample of eight well-known Dutch DOLs (for a detailed list, see Table 1) and those ordinary users who started following them during the period of analysis. Although these DOLs are not a representative nor comprehensive sample of all DOLs, they are among the most visible ones in the Dutch Twittersphere, and they are very suitable to conduct a proof-of-concept analysis to validate the theoretical concept and expectations put forward in this article. Future research should address the conditions under which the findings presented here extend to a larger and more comprehensive sample of DOLs. Despite this limitation, we believe the approach and analysis presented here contribute to building a better understanding of the actions of these types of opinion leaders and how they shape conversations on SM platforms.

Table 1. List of the eight DOLs we study.

Name	Twitter handle	Number of followers March 2, 2021	Number of followers May 5, 2021	Number of new followers analyzed (H1–H3)
Maurice de Hond	@mauricedehond	118,237	127,404 (+7.7%)	2,558
Wierd Duk	@wierdduk	84,617	90,403 (+6.8%)	1,101
Lange Frans	@langefrans	70,744	72,021 (+1.8%)	235
Robert Jensen	@robertjensen	52,686	56,462 (+7.1%)	588
Sietske Bergsma	@sbergsma	31,224	35,165 (+12.6%)	321
Café Weltschmerz	@cafeweltschmerz	17,062	17,754 (+4%)	32
Blck Bx	@blckbxnews	16,141	22,474 (+39.2%)	382
Isa Kriens	@isakriens	11,658	12,931 (+10.9%)	632
Total	—	—	32,245 (13,337 unique)	5,574

On the first day (March 2, 2021), we pulled the list of followers for each of these DOLs. We only include followers that have sent at least one tweet before to enable a comparison before and after these Twitter users started to follow a DOL. Then, every day (until May 4, 2021), we pulled the following additional information: the messages sent by the DOLs that day, the list of users who started following a given DOL that day, (up to) the last 3,200 messages sent by these new followers (to gather information about their posting behavior before following the particular DOL), and the messages posted that day by the new followers detected in previous days (to gather information about their posting behavior after they started following a particular DOL).

We use the collected data for two main purposes. First, we manually code the messages posted by the DOLs themselves for whether they contain fake news, disinformation, conspiracy, and/or rumors (non-mutually exclusive categories). The goal is to assess our claim that these DOLs engage in the dissemination of different types of deceitful content. As shown in Table 2, we rely on existing and validated definitions when coding for these four types of deceitful messages (see Part B in the Supplementary File for the codebook). For each DOL, 10 tweets were coded by two authors, resulting in 80 annotated tweets in total, leading to intercoder reliability values using Krippendorff's alpha of 0.99 for fake news, 0.98 for disinformation, 0.97 for conspiracy, and 0.93 for rumors.

Then, to test potential behavioral changes, we count the number of political (H1), uncivil (H2), and affectively polarized (H3) tweets that new followers posted during the days before versus the days after they started following the first DOL in our sample. We use two time windows for this before/after analysis, 15 and 30 days, to assess the robustness of the findings to this subjective cut-off. We collected data from 13,377 unique new followers for the DOLs in our sample. For clear identification, when testing our hypotheses, we will restrict our sample to (a) users who started following one of the DOLs after March 2, 2021 (for the previous followers, we do not know exactly the date they started following the DOL), (b) users for which we have collected their messages for the entire before and after time windows, and (c) users who did not stop following the followed DOL during data collection (a total of 3,451 users started following one of the eight DOLs under analysis, but stopped following them before the end of data collection). Our final analytical sample includes a total of 5,574 followers (see Table 1) who sent a total of 731,371 tweets during the 30 days prior/after combined.

To count the number of political, uncivil, and affectively polarized tweets, we trained three machine-learning classifiers. First, we annotated 2,896, 5,242, and 855 for whether they were uncivil, political, and affectively polarized, respectively (binary categories). Table 4 provides an overview of the annotated messages per classifier. Messages were coded as uncivil if they were insulting, harassing, very dismissive towards others, racist or against a minority group, misogynist, or when they contain curse words (Davidson et al., 2017; Theocharis et al., 2016). Messages were coded as political if (a) a political party or organization was mentioned and/or (b) if messages touched on relevant policy issues. Finally, messages were coded as being affectively polarized if users showed dislike towards an opposing group (by naming them, tagging them, or mentioning them), such as a politician, political party, or societal group (e.g., conservatives/liberals, immigrants; see Part C in the Supplementary File for the codebook). One hundred Tweets were coded by two authors, leading to intercoder reliability values using Krippendorff's alpha of 0.86 for political tweets, 0.85 for uncivil language, and 0.87 for affectively polarizing language.

Since uncivil, political, and affectively polarized tweets are rare, to have as many true positives in our annotated set as possible, we used random sampling as well as active learning when selecting the cases to be annotated (Miller et al., 2020). Hence, the number of true positives in our annotated dataset is not really a reflection of the prevalence of these quantities in the overall dataset. Table 3 shows examples of the types of messages coded as political, uncivil, and affectively polarized.

Then we used the full corpus of annotated data to fine-tune three times the same transformer model (the Dutch version of BERT [de Vries et al., 2019]—bert-base-dutch-cased), one for each of the three (political, uncivil, and affectively polarized tweets) classifiers.

Table 2. Definitions of deceitful content used for coding.

Fake news	Has a journalistic format but is low in facticity (Egelhofer & Lecheler, 2019)
Disinformation	False information that is purposely spread to deceive people, seeking to amplify social divisions and distrust (Bennett & Livingston, 2018; McKay & Tenove, 2021)
Conspiracy	Efforts to explain events, practices, or secret plots that consist of two or more powerful actors acting in secret for their benefit and working towards a malevolent or unlawful goal against the common good (Douglas et al., 2019; Sunstein & Vermeule, 2009)
Rumor	Circulating information whose veracity status is yet to be verified at the time of spreading (DiFonzo & Bordia, 2007; Friggeri et al., 2014)

Table 3. Examples of political, uncivil, and affectively polarized tweets (translated from Dutch).

Political tweet	According to the left-wing opposition parties, the deal does not go far enough, while the PVV believes that the cabinet has caved in.
Uncivil tweet	@DDStandard she's ugly. she's stupid…she's not adding anything. just a hopeless nigger who also tries to shout something…and nobody listens. she will never become someone like Pim Fortuyn…Sylvana cannot even stand in his shadow
Affective polarized tweet	RT I didn't think much of the left wing voters, but voting for fucking Sigrid Al Qaq-Kaag is like selling your soul to Europe…

Deep transformer models such as BERT have been shown to improve machine text classification in many domains, including political and communication science (Terechshenko et al., 2020). In each case, we used 20% of the annotated data to create a completely untouched validation set. Then, we split 70/30 of the remaining data into a train and test set. We used the train set to estimate model fit and update the model weights at each training iteration and the test split to assess out-of-sample performance and to decide when to stop training the model further. We stopped the training when the test loss did not improve for three complete iterations. We trained each model three times, using a different train/test split each time (three-fold cross-validation). Finally, we assessed out-of-sample accuracy on the untouched validation set (which remained constant across the three folds).

In Table 4, we report the performance of each model based on this three-fold cross-validation conducted on the validation set. The uncivil and political classifiers perform very well: Overall accuracy, as well as precision and recall, are very high; and precision and recall are very similar, indicating that in the rare cases in which a classifier makes the wrong prediction, it is equally likely to miss-classify messages that are (vs. are not) uncivil/political. The performance of the affective polarization classifier is slightly lower—high accuracy (83%) but slightly lower levels of precision (65%) and recall (71%)—but the classifier is highly balanced (similar levels of precision and recall). We have no reason to believe that there is any systematic error for any of the classifiers. So, any remaining noise would mean that we are conducting conservative tests of our hypotheses.

Finally, we use these classifiers to predict whether the rest of the unlabeled messages posted by the new 5,574 DOL followers are political, uncivil, and affectively polarized, and to count the number of political/uncivil/polarizing tweets sent the 30 days before and the 30 days after starting to follow the first DOL in our sample.

4. Results

We begin by assessing whether DOLs indeed engaged in the dissemination of many types of deceitful content (e.g., fake news, disinformation, conspiracies, and rumors). We then move to test our hypotheses regarding the behavior of new followers.

In Figure 1, we study the distribution of the deceitful content that was spread by each DOL during the period of analysis. In line with our theoretical framework, the figure illustrates that all DOLs engage (to some extent) in the dissemination of all types of deceitful content under scrutiny, from fake news to conspiracies and rumors. For example, except for Robert Jensen, the remaining DOLs posted at least one message containing each of the deceitful typologies under study. Although sometimes they have a clearly preferred deceitful category (e.g., 45% of Maurice de Hond's tweets spread fake news, and 54.8% of Blck Bx's messages promoted disinformation), they also engage in the spread of other kinds of deceitful content quite often (e.g., 16% and 9% of Maurice de Hond's tweets contained disinformation and conspiracies, respectively; and 12.5% and 7.7% of Blck Bx's messages had conspiracies and rumors in them). These results align with our argument that the main goal of these actors is to inject toxicity into online environments and that each type of deceitful content is simply one of many tools in the toolbelt of DOLs. In addition, the results emphasize that a user-centric (rather than, or in combination with, a content-centric and compartmentalized) analysis is needed to have a clearer understanding of the spread of deceitful content on SM and its effects.

To test H1, H2, and H3, we turn to the set of new followers for which we had collected enough information

Table 4. Three-fold cross-validated performance of three BERT classifiers predicting binary outcomes: Political, uncivil, and affectively polarized tweets.

Classifier	N annotated	True positives	Accuracy	Precision	Recall
Political	5,242	59%	86%	94%	86%
Uncivil	2,896	39%	86%	83%	80%
Affectively Polarized	855	35%	83%	65%	71%

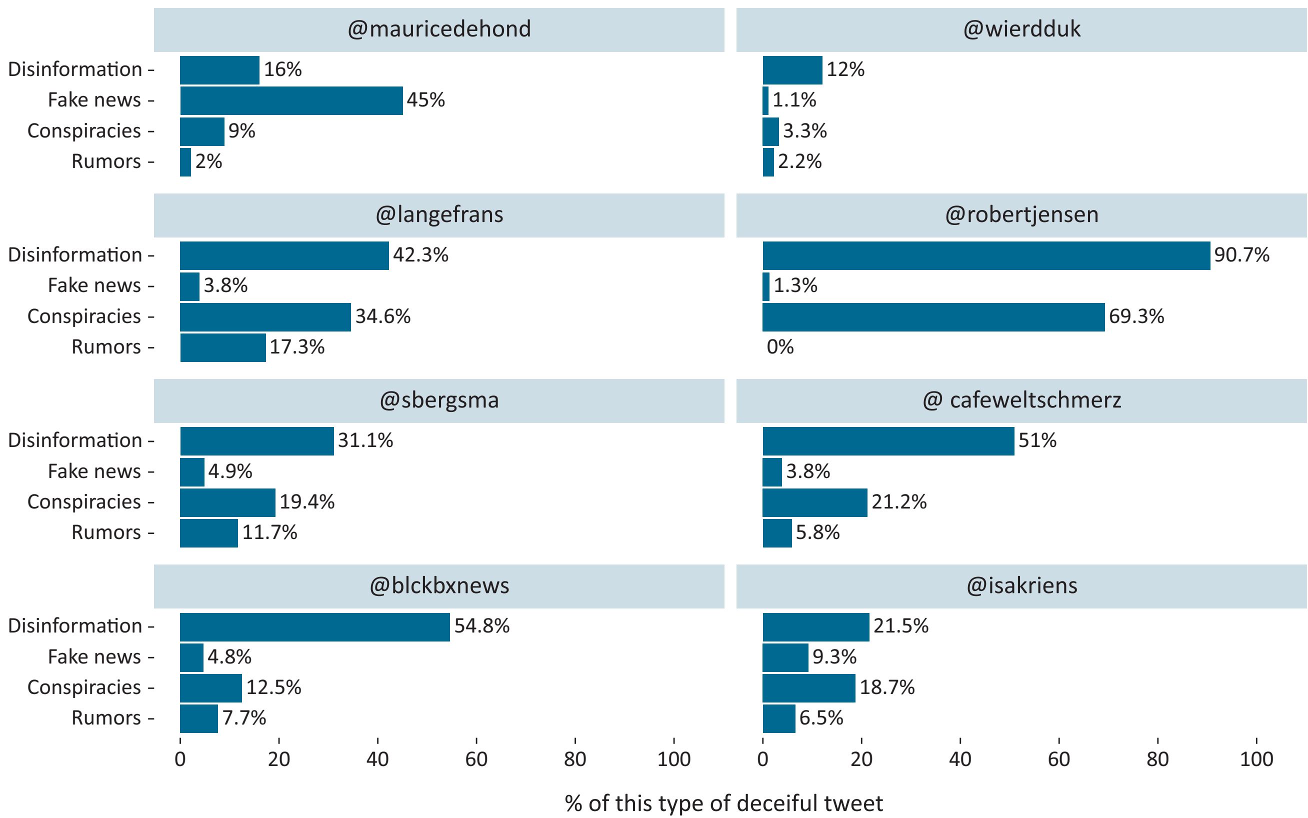

Figure 1. The percentage of tweets sent by the DOLs under analysis that contain different types of deceitful content.

to explore a potential change in behavior after following the first DOL in our sample (N = 5,574). In Figure 2, we show the average number of political, uncivil, and affectively polarized tweets these users sent during the 30 days before (vs. after) following the first DOL. We see stark differences across the board. The users were more politically engaged (sending 36.2 political tweets in the 30 days after vs. 19.3 political tweets in the 30 days prior), more uncivil (9 vs. 4.3 uncivil tweets), and affectively polarized (12.3 vs. 6.6 polarizing tweets).

Given that we collected the data during an election period, we wanted to control for whether a user started following a DOL before the election day (as users may have been more likely to discuss politics during the *after* time window). Hence, we created the variable Campaign Post Days, which accounts for the number of *post* 15/30 days that overlapped with the electoral campaign (so the number of days between the day a user started following the first DOL and election day, March 17). This variable is 0 for those who started following a deceitful opinion leader after March 17. As specified in Model 1, for a clearer test of our hypotheses we use linear models predicting the difference (Y_{post} – Y_{pre}) for three outcomes of interest (number of uncivil, affectively polarized, and political tweets) as a function of the mentioned control variable Campaign Post Days. For each of these linear models, the intercept parameter (α) provides information about the average difference in messaging behavior between the post and pre-difference after accounting for the control variable.

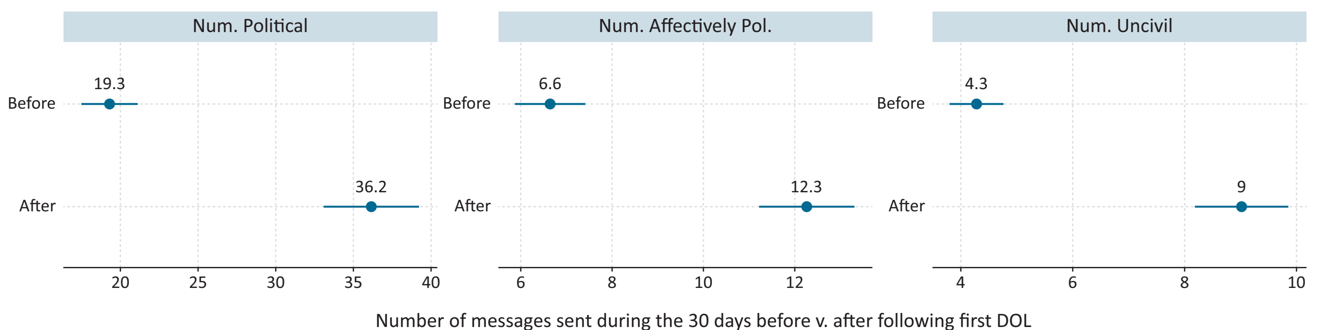

Figure 2. Average number of political, uncivil, and affectively polarized tweets (plus 95% confidence interval) sent during the 30 days before (vs. after) following the first DOL.

Model 1, the model specification used to test H1, H2, and H3, is as follows:

$$(Y_{post} - Y_{pre}) = \alpha + \beta_{CampaigPosDays} + \varepsilon$$

In Figure 3, we report the α coefficient for several linear models. For each of the four outcomes of interest, we ran six models with the same specification (i.e., Model 1), where we varied the time window to calculate the post/pre periods (15- and 30-day windows) and the number of DOLs the user followed within the 15/30 days after following the first opinion leader. In the first column (1 DOL), we include all the users in our sample (N = 5,574/3,891), and in the other columns we estimate the models using only those users who followed at least a second DOL (2 DOLs) within the next 15/30 days (N = 1,336/1,014), and at least a third DOL (3 DOLs; N = 555/421); i.e., each analysis includes the number of unique Twitter users that meet the criteria. These variations allow us to disentangle differential effects across time (whether we observe stronger effects when comparing 15 vs. 30 days), and across different levels of engagement (e.g., users who decided to follow more than one of the DOLs in our sample).

We find strong support for our three hypotheses. Across the board, we see an increase in the number of political, uncivil, and affectively polarized tweets. All estimates presented in Figure 3 are statistically significant at the conventional 0.05 level. We observe the mildest effects among those who only followed one of the DOLs in our sample. But even among those, we observe a substantial change in behavior, particularly when we compare the behavior during the 30 days after (vs. before) following the DOL. On average, these users sent 27.4 more political messages, 8.8 more affectively polarized messages, and 7.9 more uncivil ones. We observe the strongest effects among those who followed a second and a third DOL during the 30 days after following the first DOL in our sample (N = 421). On average, they radically sent more political (+108.4), affectively polarized (+38.2), and uncivil (+29.1) messages. These findings are not driven by the new followers of one particular DOL, but reflect a general pattern observed across the followers of the different DOLs in our sample (see Appendix D in the Supplementary File). In addition, this behavior change cannot be simply explained by these users retweeting messages originally posted by the DOLs they started following (see Appendix A in the Supplementary File). On average (95% confidence intervals included), only 0.8% (0.6–1%) of the political tweets, 0.5% (0.3–0.7%) of the uncivil messages, and 0.5% (0.3–0.7%) of the affectively polarized tweets they sent during the 30 days after following the first DOLs are retweets of that DOL.

For a more detailed picture, in Figure 4, we explore the functional form of these effects. The figure shows the average number of political, uncivil, and affectively polarized tweets (+95% confidence intervals) the users in our sample sent each of the 30 days before and after following the first DOL. Figure 4 is standardized, so the exact date of day 0 differs across users, depending on when they started following the DOL. We observed a slight upper trend right before they started following the DOL. This indicates that at least some users already started shaping their behavior before day 0. This could be because they may have already been exposed to some tweets from these DOLs via retweets from their networks, or some factor motivated them to change their behavior and potentially seek these kinds of opinion

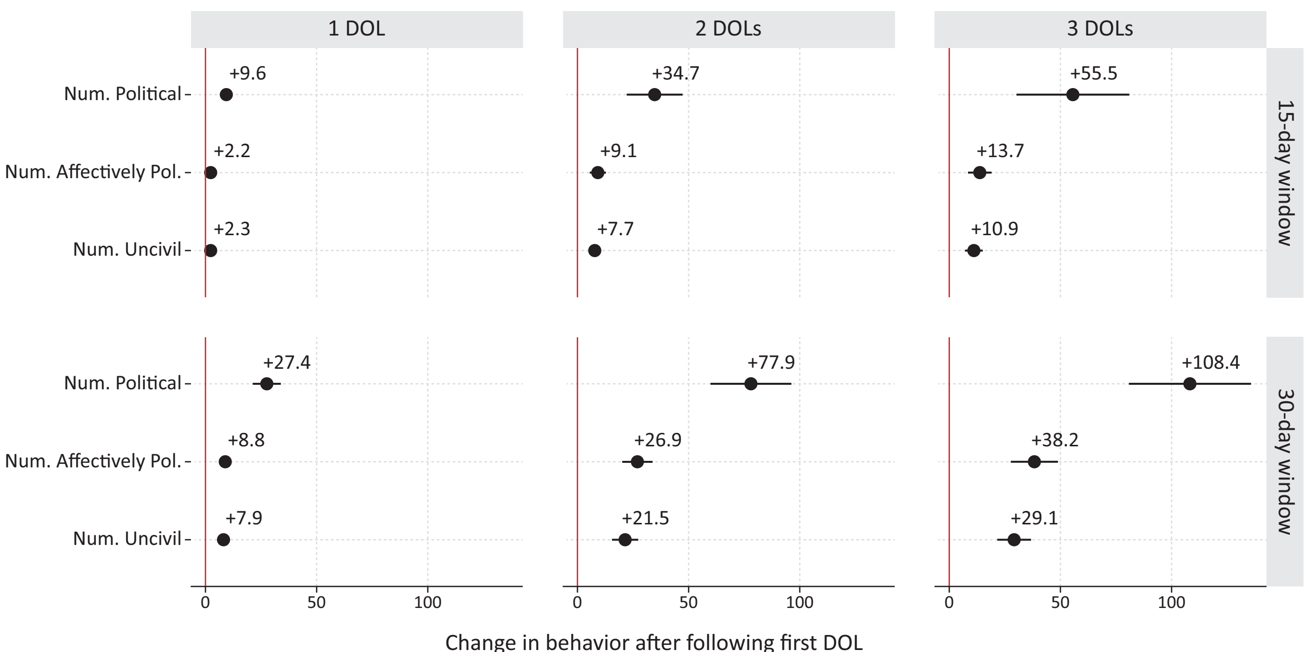

Figure 3. Coefficients (+95% confidence intervals) from linear models estimating a change in behavior after following one, two, and three DOLs.

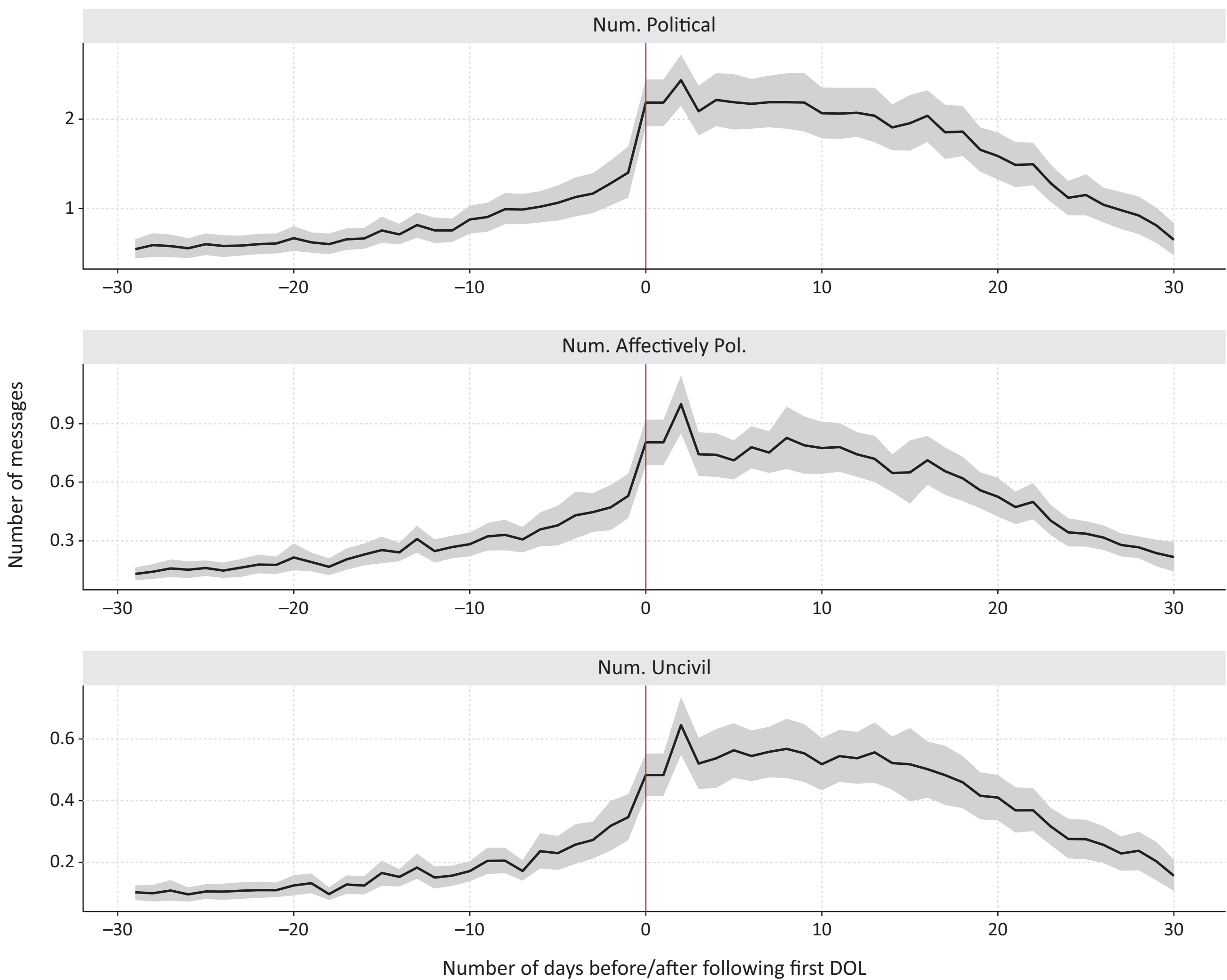

Figure 4. Average number of political, uncivil, and affectively polarized tweets sent by followers of DOLs, during the 30 days before and after following the first DOL (N = 3,891).

leaders. Then, we observe a clear jump at the moment the users started following the first DOL. The number of overall tweets and the uncivil, affectively polarized, and political ones remained high for about 15 to 20 days. After that period, the behavior of the users gradually reverted to their levels of activity before following the DOL. The patterns described in Figure 4 clearly point to these DOLs playing a crucial role in the radicalization of online environments. Independently of what motivated these users to start following these DOLs—whether it was a very intentional decision or because of incidental exposure via retweets from one's networks—we observe stark and substantive changes in behavior that contribute to increasing levels of toxicity and incivility on the SM platform.

5. Conclusion and Discussion

This article tackles three shortcomings of existing literature studying the dissemination of deceitful content. First, existing literature is very compartmentalized, as it mostly focuses on one type of deceitful content (e.g., conspiracies, fake news, or misleading information). We show that salient SM accounts engage in the spread of all sorts of deceitful content throughout. Each type of deceitful content that they disseminate is just one of many tools in their toolbelt. Second, we lack an overarching approach that puts these influential SM users at the center instead of the content that they spread. We do so by putting forward a new theoretical concept: "deceitful opinion leaders." Third, this study contributes to the understanding of the individual-level effects that these types of deceiving messages have on other SM users. We show that after following a DOL on Twitter, significant behavioral changes start to occur amongst their followers: users send more political, uncivil, and affectively polarized messages. For example, on average, the analyzed users sent around 28 political tweets, 8 uncivil tweets, and 9 affectively polarized tweets more during the 30 days after following a DOL, compared to the 30 days prior. These behavioral changes seem to gradually revert to their levels of activity before following the

DOL. Although at the individual level these behavioral changes do not last long, at the aggregate level these effects have a substantive impact: DOLs gather new followers every day, meaning that these behavioral effects are constantly occurring, having a longer-lasting effect on the behavior and norms of conversation on Twitter.

Although this article adds important results to existing literature, it is not without limitations. This article provides a first aim in studying the effects of DOLs on SM platforms. There are other influential DOLs who were not included in this research. Moreover, all the DOLs in this study are Dutch. Hence, this study only focuses on the Dutch SM landscape. Furthermore, this study only considers Twitter, while DOLs are active on many platforms. To assess the generalizability of the effects that DOLs have on other SM users, future studies should aim to address additional factors that influence these findings, such as platform affordances and the level of radicalization of a platform. We expect the work presented here to inspire future work focusing on a more comprehensive and representative sample of DOLs from different contexts on different platforms, to provide further insights into the conditions under which these opinion leaders shape our online environments. Despite these limitations, this research finds valid and important results that show significant individual-level effects from following DOLs who engage in the spread of deceitful content online. Even though research finds that only a small proportion of SM users spread deceitful content per se (e.g., Guess et al., 2019), the spread of deceitful content via SM leads to substantial effects on other users on the platform.

The results of this study provide a first look into the distribution of the spread of deceitful content by DOLs and the individual-level effects that DOLs have on their followers. Importantly, this study adds to the empirical evidence of the effects of deceitful content on SM users. The findings of this study add to existing literary knowledge of the consequences of deceitful content in online environments. In addition, the results of this study provide empirical evidence to the societal debate on whether these influential SM users should be removed to maintain a healthy forum for public debate. Removing DOLs from Twitter would reduce toxicity on the platform. However, doing so might have negative effects if DOLs move on to other platforms to spread their deceiving content. This might result in higher levels of radicalization and polarization. Especially on Telegram, which is known to have a high number of users that support conspiracy theories. Furthermore, these findings underline that a small group of DOLs change the norms of conversation on SM platforms. Hence, this accentuates the need for future research to study the literary concept of DOLs.

Acknowledgments

This research is funded by an NWA Small Project Grant ("When Words Do the Trick: The Effect of Rhetorical Strategy on the Perceived Legitimacy of Political Decision-Making") awarded to Dr. Mariken van der Velden and Dr. Andreu Casas Salleras. We thank the members of the Political Communication Group at the Free University Amsterdam, and the participants of the workshop "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" and the panel "How the Type of Media Use Affects Belief in Conspiracy Theories and Misinformation" at ICA'22 for their thoughtful comments and suggestions on early versions of this manuscript.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Balkin, J. M. (2017). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UCDL Review*, *51*, 1149–1210.

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, *23*(1), 76–91. https://doi.org/10.1093/pan/mpu011

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, *26*(10), 1531–1542. https://doi.org/10.1177/0956797615594620

Baum, M. A. (2002). Sex, lies, and war: How soft news brings foreign policy to the inattentive public. *American Political Science Review*, *96*(1), 91–109. https://doi.org/10.1017/S0003055402004252

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bergmann, E. (2020). Populism and the politics of misinformation. *Safundi*, *21*(3), 251–265. https://doi.org/10.1080/17533171.2020.1783086

Bergström, A., & Jervelycke Belfrage, M. (2018). News in social media: Incidental consumption and the role of opinion leaders. *Digital Journalism*, *6*(5), 583–598. https://doi.org/10.1080/21670811.2018.1423625

Berr, J. (2019, June 25). Conspiracy theorist Alex Jones evades crackdown on social media sites. *Forbes*. https://www.forbes.com/sites/jonathanberr/2019/06/24/conspiracy-theorist-alex-jones-evades-crackdown-on-social-media-sites/?sh=1809755a4f87

Carlson, T. N. (2019). Through the grapevine: Infor-

mational consequences of interpersonal political communication. *American Political Science Review*, *113*(2), 325–339. https://doi.org/10.1017/S000305541900008X

Chadwick, A., & Stanyer, J. (2022). Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: Toward a holistic framework. *Communication Theory*, *32*(1), 1–24. https://doi.org/10.1093/ct/qtab019

Choi, S. (2015). The two-step flow of communication in Twitter-based public forums. *Social Science Computer Review*, *33*(6), 696–711. https://doi.org/10.1177/0894439314556599

Coaston, J. (2018, August 6). Alex Jones banned from YouTube, Facebook, and Apple, explained. *Vox*. https://www.vox.com/2018/8/6/17655658/alex-jones-facebook-youtube-conspiracy-theories

Davidson, T., Macy, M., Warmsley, D., & Weber, I. (2017). *Automated hate speech detection and the problem of offensive language*. arXiv. https://arxiv.org/abs/1703.04009

de Vries, W., van Cranenburgh, A., Bisazza, A., Caselli, T., van Noord, G., & Nissim, M. (2019). *BERTje: A Dutch BERT model*. arXiv. https://doi.org/10.48550/ARXIV.1912.09582

DiFonzo, N., & Bordia, P. (2007). Rumor, gossip and urban legends. *Diogenes*, *54*(1), 19–35. https://doi.org/10.1177/0392192107073433

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding conspiracy theories. *Political Psychology*, *40*(S1), 3–35. https://doi.org/10.1111/pops.12568

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, *43*(2), 97–116. https://doi.org/10.1080/23808985.2019.1602782

Friggeri, A., Adamic, L., Eckles, D., & Cheng, J. (2014). Rumor cascades. In E. Adar & P. Resnick (Eds.), *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 101–110). Association for the Advancement of Artificial Intelligence.

Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics*, *12*(2), 167–185. https://doi.org/10.1080/19331681.2014.997416

Groshek, J., & Koc-Michalska, K. (2017). Helping populism win? Social media use, filter bubbles, and support for populist presidential candidates in the 2016 US election campaign. *Information, Communication & Society*, *20*(9), 1389–1407. https://doi.org/10.1080/1369118X.2017.1329334

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, *5*(1), Article eaau4586. https://doi.org/10.1126/sciadv.aau4586

Harteveld, E. (2021). Fragmented foes: Affective polarization in the multiparty context of the Netherlands. *Electoral Studies*, *71*, Article 102332. https://doi.org/10.1016/j.electstud.2021.102332

Haselton, T. (2019, May 3). Alex Jones was banned from Facebook, but an hour later he was back on Facebook livestreaming. *CNBC*. https://www.cnbc.com/2019/05/02/alex-jones-banned-from-facebook-but-hes-already-back.html

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology. *Public Opinion Quarterly*, *76*(3), 405–431. https://doi.org/10.1093/poq/nfs038

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communications*. Free Press.

Kim, J. W., Guess, A., Nyhan, B., & Reifler, J. (2021). The distorting prism of social media: How self-selection and exposure to incivility fuel online comment toxicity. *Journal of Communication*, *71*(6), 922–946. https://doi.org/10.1093/joc/jqab034

Kim, Y., Chen, H.-T., & Gil de Zúñiga, H. (2013). Stumbling upon news on the Internet: Effects of incidental news exposure and relative entertainment use on political engagement. *Computers in Human Behavior*, *29*(6), 2607–2614. https://doi.org/10.1016/j.chb.2013.06.005

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

McKay, S., & Tenove, C. (2021). Disinformation as a threat to deliberative democracy. *Political Research Quarterly*, *74*(3), 703–717. https://doi.org/10.1177/1065912920938143

Miller, B., Linder, F., & Mebane, W. R. (2020). Active learning approaches for labeling text: Review and assessment of the performance of active learning approaches. *Political Analysis*, *28*(4), 532–551. https://doi.org/10.1017/pan.2020.4

Paul, K. (2019, May 3). Facebook bans Alex Jones, other extremist figures. *Reuters*. https://www.reuters.com/article/us-facebook-extremists-usa-idUSKCN1S82D7

Popan, J. R., Coursey, L., Acosta, J., & Kenworthy, J. (2019). Testing the effects of incivility during internet political discussion on perceptions of rational argument and evaluations of a political outgroup. *Computers in Human Behavior*, *96*, 123–132. https://doi.org/10.1016/j.chb.2019.02.017

Rosdorff, M. (2018, August 8). Complotdenker Alex

Jones verbannen van bijna alle sociale media [Conspiracy theorist Alex Jones banned from almost all social media]. *EenVandaag*. https://eenvandaag.avrotros.nl/item/complotdenker-alex-jones-verbannen-van-bijna-alle-sociale-media

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36. https://doi.org/10.1145/3137597.3137600

Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, *34*(3), 150–160. https://doi.org/10.1177/0266382117722446

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, *30*(3), 341–366. https://doi.org/10.1007/s11109-007-9050-9

Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, *7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Sunstein, C. R., & Vermeule, A. (2009). Conspiracy theories: Causes and cures. *Journal of Political Philosophy*, *17*(2), 202–227. https://doi.org/10.1111/j.1467-9760.2008.00325.x

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In M. J. Hatch & M. Schultz (Eds.), *Organizational identity: A reader* (pp. 55–65). Oxford University Press.

Terechshenko, Z., Linder, F., Padmakumar, V., Liu, F., Nagler, J., Tucker, J. A., & Bonneau, R. (2020). *A comparison of methods in political science text classification: Transfer learning language models for politics*. SSRN. http://doi.org/10.2139/ssrn.3724644

Theocharis, Y., Barberá, P., Fazekas, Z., Popa, S. A., & Parnet, O. (2016). A bad workman blames his tweets: The consequences of citizens' uncivil Twitter use when interacting with party candidates. *Journal of Communication*, *66*(6), 1007–1031. https://doi.org/10.1111/jcom.12259

Van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., de Vreese, C., Matthes, J., Hopmann, D., Salgado, S., Hubé, N., Stępińska, A., Papathanassopoulos, S., Berganza, R., Legnante, G., Reinemann, C., Sheafer, T., & Stanyer, J. (2017). Political communication in a high-choice media environment: A challenge for democracy? *Annals of the International Communication Association*, *41*(1), 3–27. https://doi.org/10.1080/23808985.2017.1288551

Weeks, B. E., & Gil de Zúñiga, H. (2021). What's next? Six observations for the future of political misinformation research. *American Behavioral Scientist*, *65*(2), 277–289. https://doi.org/10.1177/0002764219878236

Weeks, B. E., Lane, D. S., Kim, D. H., Lee, S. S., & Kwak, N. (2017). Incidental exposure, selective exposure, and political information sharing: Integrating online exposure patterns and expression on social media. *Journal of Computer-Mediated Communication*, *22*(6), 363–379. https://doi.org/10.1111/jcc4.12199

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021). Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Political Communication*, *38*(1/2), 98–139. https://doi.org/10.1080/10584609.2020.1785067

YouTube verdedigt verwijderen account Lange Frans: "Richtlijnen meermaals geschonden" [YouTube defends the removal of Lange Frans account: "Guidelines have been violated several times"]. (2020, October 21). *AD*. https://www.ad.nl/show/youtube-verdedigt-verwijderen-account-lange-frans-richtlijnen-meermaals-geschonden~a5606d7d

Zaller, J. (1992). *The nature and origins of mass opinion*. Cambridge University Press.

About the Authors

Puck Guldemond is a PhD candidate at the Strategic Communication Group at Wageningen University & Research. Her research interests include data-driven campaigning, political influence and behavior on social media, computational social sciences, and populist rhetoric in political communication.

Andreu Casas Salleras is an assistant professor at the Department of Communication Science at the Vrije Universiteit Amsterdam and a faculty associate at the Center for Social Media and Politics at New York University. His research interests encompass the areas of political communication, public policy processes, and computational social sciences. He is particularly interested in how social movements and interest groups influence the political agenda and the decision-making process in the current media environment.

Mariken van der Velden is an associate professor of political communication at the Vrije Universiteit Amsterdam. She was a postdoctoral researcher at the Institute of Political Science at the University of Zurich and visiting researcher at the Departments of Political Science at the University of North Carolina, Chapel Hill, and the University of Oxford. Her research interest comprises the areas of political communication, political behavior, and computational social science.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 349–360
https://doi.org/10.17645/pag.v10i4.5493

Article

Who Are the "Dark" Politicians? Insights From Self-Reports of German State Parliament Candidates

Jürgen Maier [1,]*, Mona Dian [1], and Corinna Oschatz [2]

[1] Department of Political Science, University of Koblenz-Landau, Germany
[2] Amsterdam School of Communication Research, University of Amsterdam, The Netherlands

* Corresponding author (maierj@uni-landau.de)

Submitted: 4 March 2022 | Accepted: 7 September 2022 | Published: 30 December 2022

Abstract
A growing body of studies is focusing on politicians' personalities, as the personality of political elites has been shown to affect their behavior. Whereas most research uses the big five framework or HEXACO, only a few studies have been able to capture more "aversive," "dark"—yet non-pathological—personality traits of politicians. However, these studies refer to top politicians; information on the distribution and the correlates of dark personality traits in the broad mass of politicians is still lacking. Moreover, information on dark personality traits in politicians is usually based on expert ratings; data using self-placement is missing. Based on data from six surveys with candidates running for German state elections in 2021 and 2022 (N[pooled data set] = 1,632), we, to the best of our knowledge, offer, for the first time, insights into politicians' self-reported socially aversive personality traits. "Dark" personality traits are measured by the political elites aversive personality scale (PEAPS). Results show that German politicians exhibit moderate levels of aversive personality traits. In addition, the extent of candidates' dark personalities is strongly negatively correlated with honesty–humility, agreeableness vs. anger, and extraversion, while associations with other basic personality traits are much weaker or insignificant. We also find that younger, more right-leaning, and more ideologically extreme candidates report higher levels of aversive personality.

Keywords
aversive personality; candidate survey; dark personality; Germany; self-reports

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

The media often depict top politicians based not only on their actions but also on their personalities. For example, the media attributed narcissism (among other traits) to former US President Donald Trump (Bannon, 2020), former German Chancellor Angela Merkel has been described as conscientious and patient (Bildt, 2018), and current Russian President Vladimir Putin has been seen—even before Russia's attack on Ukraine—as deceitful and aggressive ("Vladimir Putin: Russia's action man president," 2021). In addition, several scientific studies have analyzed the personality of (top) politicians (e.g., Nai, 2019a; Nai & Maier, 2018; Rice et al., 2021; Visser et al., 2017); their results show that voters not only have a choice between different policy programs when it comes to elections but also between different personalities of political leaders. Because personality has an impact on their performance (for instance, policy successes, relationships with the legislature, use of executive orders, and the likelihood of unethical behavior; see, e.g., Lilienfeld et al., 2012; Rubenzer & Faschingbauer,

2004; Rubenzer et al., 2000; Watts et al., 2013), analyzing the personality of political elites is not only of academic interest but also of practical relevance for better understanding the outcomes of political systems.

Despite all efforts to measure politicians' personalities, current research has some limitations. First, most studies focus on those already in office (e.g., members of parliament). Studies on candidates running for office are less common. However, candidate studies can be used to determine the pool of personnel from which voters can select their representatives. Only by comparing successful candidates (i.e., members of parliament) with unsuccessful candidates we gain insights into the factors that promote electoral success and the contribution of personality (see, e.g., Joly et al., 2019; Scott & Medeiros, 2020). Second, most research has focused on politicians' basic personality traits, usually measured via the big five framework (e.g., McCrae & John, 1992) and, to a lesser extent, the HEXACO model of personality (e.g., Ashton & Lee, 2007; Best, 2011; Hanania, 2017; Maier & Nai, 2021; Nai, 2019a; Ramey et al., 2019; Rice et al., 2021; Schumacher & Zettler, 2019; Visser et al., 2017). However, research suggests that individuals also differ in socially aversive—yet non-pathological—personality traits that are responsible for behavior violating generally accepted ethical, moral, and social norms. Therefore, analyzing the "dark" personality of politicians might help to better understand more recent developments in political communication and political behavior, which are often considered dysfunctional for the effectiveness of the political discourse and, more generally, a threat to the cohesion of society—e.g., negativity, incivility, populism, spreading of fake news, and involvement in scandals. However, the number of available studies on politicians' aversive personalities is still very limited (see Lilienfeld et al., 2012; Nai, 2019a, 2019b, 2022; Nai & Martínez i Coma, 2019; Nai et al., 2019; Nai & Toros, 2020; Simonton, 1988). Third, unlike the assessment of basic personality traits, where self-reports are not uncommon, the measurement of the "dark" personality of politicians is still limited to the assessments of experts (e.g., Lilienfeld et al., 2012; Nai, 2019a) and voters (Nai & Maier, 2021a). The reasons for this seem obvious: On the one hand, available inventories to measure aversive personality traits are often quite long; politicians are unlikely to spend much time (if any; this is especially true for top politicians who are unlikely to participate at all) in answering questions related to the "dark" sides of their personality. On the other hand, the available standard inventories of aversive personality often contain relatively difficult questions that are hard to present to a politician and thus pose a threat to the successful completion of the survey. Hence, to have any chance of collecting self-assessments of "dark" personality traits, we need a brief and inoffensive inventory. The current lack of such a self-assessment has important consequences for research. In fact, it limits our understanding to a few very prominent politicians of larger established parties.

This study helps to fill the identified research gaps to some extent. Based on six German candidate surveys, we provide insights into the distribution and predictors of candidates' self-reported levels of aversive ("dark") personality traits. To do so, we use a short scale designed to measure politicians' self-assessed aversive personalities. This allows us to make statements that go beyond a limited selection of top and prominent politicians and also include politicians from smaller parties. The results suggest that German candidates have moderate levels of aversive personality traits and that the extent of candidates' aversive personalities can be predicted by candidates' social and political characteristics. Selected characteristics (age, ideology, and extremism) remain significant predictors of aversive personality even after controlling for basic personality traits.

2. Dark Personality: Measurement and Correlates

2.1. Measurement

There is neither a universally accepted model of aversive personality nor agreement on how to measure it. The most prominent framework in this regard is the so-called dark triad of personality, consisting of three related yet (allegedly) distinctive personality traits: narcissism, psychopathy, and Machiavellianism (Paulhus & Williams, 2002). In a nutshell, psychopathy includes "impulsivity, reckless risk-taking, and very shallow empathy toward other people" (Lyons, 2019, p. 2). Narcissism is the belief that one has "superior abilities in comparison to other people" (Lyons, 2019, p. 3), coupled with seeking "attention and admiration" (agentic narcissism) "while devaluing others" (antagonistic narcissism; Rauthmann, 2012, p. 487). Machiavellianism is "the flexible, chameleon-like use of strategies from defection to cooperation to suit the demands of the situation, with the ultimate aim of gaining benefits for the self" (Lyons, 2019, p. 2). The dark triad has been successfully used to measure the aversive personality traits of political candidates based on expert ratings (e.g., Lilienfeld et al., 2012; Nai, 2019a) and voter ratings (Nai & Maier, 2021a).

The concept of the dark triad is not unchallenged. For example, some scientists claim that the dark triad should be expanded to include sadism, a trait that measures the reward "of inflicting unnecessary pain on others" (Lyons, 2019, p. 35). This four-domain concept is known as the dark tetrad (e.g., Chabrol et al., 2009). With respect to politicians, "everyday sadism" (Buckels et al., 2013), i.e., non-pathological "behaviors that [are] not too extreme or illegal" (Lyons, 2019, p. 36), could be relevant but have not yet been studied. Furthermore, other scholars claim that aversive personality traits, such as those measured by the dark triad, are manifestations of a common *"dark" core of personality* (e.g., Moshagen et al., 2018, 2020; Schreiber & Marcus, 2020; Vize et al., 2020). In particular, the idea of a "common core" of aversive personality is intriguing for the

study of self-reported "dark" personality in politicians because its measurement is not tied to a fixed set of items but is explicitly understood as a "fluid construct...that...appears in all combinations of a sufficient number of different indicators of dark traits in a form that mirrors our conceptualization" (Moshagen et al., 2018, p. 659). Consequently, the concept of there being a dark core of personality is not limited to the dimensions proposed by the dark triad (i.e., narcissism, psychopathy, and Machiavellianism) but is also open to other socially aversive personality traits (e.g., amoralism, egoism, moral disengagement, sadism, self-centeredness, and spitefulness; see Moshagen et al., 2020).

This allows the compilation of items suitable in number and content to be presented to politicians. Maier et al. (2022) recently proposed such an instrument, the political elites aversive personality scale (PEAPS), which we will use in our study.

The few available studies on the aversive personality of politicians show that, on average, political leaders exhibit high levels of narcissism and moderate levels of psychopathy and Machiavellianism (Nai, 2019a, 2019b, 2022; Nai & Martínez i Coma, 2019; Nai & Toros, 2020). Nai (2022) calculated the mean across the three domains of the dark triad to measure the dark core of personality. Using a sample of 49 top candidates from 22 countries, he reported a mean of 2.6 on a scale from 0 (*low level of dark personality*) to 4 (*high level of dark personality*).

2.2. Correlates

There is little research on which social and political characteristics correlate with politicians' "dark" personality traits. First, Nai and Martínez i Coma (2019) have shown that while female populist candidates exhibit lower levels of narcissism, there are no gender differences in psychopathy and Machiavellianism. Nai and Maier (2020) find no correlation between politicians' gender and the level of dark personality. However, psychological research provides strong evidence that females score lower on aversive personality traits than males (e.g., Muris et al., 2017; Schmitt et al., 2017). These differences are, on the one hand, explained by biological factors—e.g., genetic dispositions or prenatal experiences (Schmitt et al., 2017). On the other hand, social factors (e.g., differences in socialization, gender roles, gender stereotypes, or the balance of power between males and females in a society) account for gender differences in dark personality traits (Schmitt et al., 2017), as the rewards and punishments for exhibiting aversive behaviors are not the same for men and women, but typically favor the former and punish the latter (Landay et al., 2019). Therefore, we expect that males show higher levels of "dark" personality than females (H1).

Second, the relationship between (politicians') aversive personality and age lacks coherence. Nai and Martínez i Coma (2019) found that younger populist leaders score somewhat higher on narcissism than older candidates. In contrast, there was no relationship between psychopathy and Machiavellianism. However, from the perspective of personality development, we can expect that aversive personality traits should be negatively correlated with age since older individuals are better able to regulate their emotions (Carstensen et al., 2003) and therefore tend to be more agreeable (Chopik & Kitayama, 2018; Roberts et al., 2006). Consistent with this increase in "psychological maturity" (Roberts et al., 2006, p. 3; as well as because people adjust their life history strategy over the course of their lives in response to perceived changes in the degree of unpredictability of their environment [Hartung et al., 2022] and that prosocial behavior becomes more beneficial as people assume more "stable" social roles [Roberts & Wood, 2006]) socially aversive personality traits have been shown to decrease with age (Hartung et al., 2022; Klimstra et al., 2020). Therefore, we expect a negative relationship between "dark" personality and age (H2).

Third, Nai and Martínez i Coma (2019) report that incumbent populist candidates show higher levels of Machiavellianism, narcissism, and psychopathy than challengers. Furthermore, there is evidence that personality traits influence the decision to run for office (Scott & Medeiros, 2020) as well as electoral success (Joly et al., 2019; Scott & Medeiros, 2020). Among aversive personality traits, Machiavellianism and narcissism are positively related to political success (Blais & Pruysers, 2017). The findings from the political sphere are confirmed by studies from the business world, which show that psychopaths are more likely to hold leadership positions or be perceived as leaders (Landay et al., 2019). The correlation between aversive personality and political success could be due to two reasons: On the one hand, individuals with dark personality traits score high on social values that are relevant for obtaining leadership positions. In particular, dark personalities consider achievement (e.g., success, ambition) and power (e.g., authority, wealth) as important goals (Kajonius et al., 2015). On the other hand, individuals with higher levels of aversive personality exhibit certain skills that are useful for success. For instance, psychopaths are described as "calm and focused in situations involving pressure or threat" (Patrick et al., 2009, p. 926). Therefore, we expect that incumbents self-report a higher level of aversive personality than challengers (H3). Furthermore, we expect that candidates scoring high on aversive personality have a higher likelihood of being elected (H4).

Fourth, there is evidence that aversive personality is positively correlated with left-right ideological placement, i.e., more conservative politicians show higher levels of "dark" personality (Nai & Maier, 2020; Nai & Martínez i Coma, 2019). This finding matches the result that conservative politicians show lower levels of agreeableness than liberals (e.g., Caprara & Vecchione, 2017, p. 224; Dietrich et al., 2012; see also Schumacher & Zettler, 2019). One explanation for the finding that the left–right placement of politicians is positively correlated

with their level of dark personality is that conservative beliefs are linked to social dominance (e.g., Van Hiel & Mervielde, 2002). Social dominance, in turn, is positively associated with psychopathy, narcissism, and Machiavellianism (Jones, 2013). Subjects scoring high on social dominance tend to have a "'social Darwinist' view of the world," i.e., "a view of the world as a ruthlessly competitive jungle in which the strong win and the weak lose" (Duckitt, 2006, p. 685). Therefore, we expect that left-leaning (right-leaning) politicians self-report lower (higher) levels of aversive personality (H5).

Finally, several studies reported a link between aversive personality and extreme political attitudes, extreme (intended) political behavior, or preference for left- or right-wing political parties among citizens (for an overview, see Pavlović & Wertag, 2021). However, the reasons for this relationship are not clear yet. Different traits considered as "dark" might explain this link, for instance, "feelings of being treated unjustly, intolerance to frustration, hostile reactions to perceived provocation and discrimination, depreciation and dehumanization of the victims" (narcissism); "detachment from conventional morality and search for power, control, and authority" (Machiavellianism); "impulsivity, low empathy and callous disregard of others" (psychopathy); or "the feeling of pleasure derived from dominance and suffering of others" (sadism; Chabrol et al., 2020, p. 158). Although to the best of our knowledge, there is no empirical evidence on the relationship between political extremism and dark personality traits in politicians, we nevertheless assume the correlation to be similar to that of the general public. Therefore, we expect politicians who self-report an extreme ideological position to exhibit higher levels of aversive personality (H6).

3. Research Design

3.1. Data

Our analyses are based on post-election surveys of candidates running in six state elections in Germany in 2021 and 2022: Baden-Wuerttemberg, Berlin, Mecklenburg-Western Pomerania, Rhineland-Palatinate, Saxony-Anhalt (all 2021), and Saarland (2022). Data were collected using a mixed mode. Data collection began the day after election day and ended two months later. Approval of the Institutional Review Board (IRB) was obtained prior to data collection (the GESIS ethics committee approved the study on 27 November 2020, Reference No. 2020–6). All candidates (including candidates running for smaller parties in the 2021 elections) were invited to participate. Candidates that provided an email address in their professional contact details online were invited to participate via an online link to our survey (Baden-Wuerttemberg: 81.4%; Berlin: 56.6%; Mecklenburg-Western Pomerania: 65.0%; Rhineland-Palatinate: 66.8%; Saxony-Anhalt: 58.5%; Saarland: 52,1%). All candidates without a publicly available email address were invited by regular mail, including a paper-and-pencil questionnaire and a return envelope. They were also provided with a personalized link in case they preferred to answer the survey online. Since many candidates in state elections can rely on campaign or office staff if they are already members of parliament, we explicitly asked candidates in the invitation letter to complete the questionnaire themselves. From the initial 3,842 candidates contacted (Baden-Wuerttemberg: 824; Berlin: 1,116; Mecklenburg-Western Pomerania: 434; Rhineland-Palatinate: 788; Saxony-Anhalt: 423; Saarland: 257), 43.5% participated in the study (minimum 35.0%, maximum 59.5%). Note that there are no significant differences in social and political profile between participants and non-participants except for incumbency (i.e., incumbents participated significantly less often in the survey; see Table C1 of Appendix C in the Supplementary File). For our analyses, we excluded 41 candidates who rushed through the (online) survey by employing the procedure to filter out speeders described by Leiner (2019). This resulted in N = 1,632 valid cases (Baden-Wuerttemberg: 473; Berlin: 382; Mecklenburg-Western Pomerania: 158; Rhineland-Palatinate: 354; Saxony-Anhalt: 151; Saarland: 114).

One-third (33.3%) of the participants were female. The average age was 45.3 years (*SD* = 13.7 years). Data for gender and age is based on the information of the state returning officer (*Landeswahlleiter*). The IRB approval covers linking candidates' survey responses with external sources; these linking possibilities were explicitly mentioned to the candidates in the informed consent. The ideology of the sample was slightly skewed to the left (*M*: 4.72/1–11, *SD*: 2.18); 12.3% of the candidates who participated in our surveys ran for the Christian Democrats (CDU), 12.2% for the Social Democrats (SPD), 12.1% for the Green Party (Bündnis 90/Die Grünen), 8.6% for the Left Party (Die Linke), 11.5% for the Liberal Party (FDP), 5.2% for the Alternative for Germany (AfD), and 38.1% for smaller parties not (yet) represented in the parliament. On average, participants took 17 minutes and 45 seconds (*SD* = 325 s) to complete the questionnaire (information only available for the online questionnaires).

3.2. Measures

3.2.1. Dependent Variable

Aversive ("dark") personality was measured by the PEAPS (Maier et al., 2022; for more information on the development of psychometric characteristics, see Appendix A in the Supplementary File). The scale is a six-item short scale developed specifically to measure the self-reported aversive personality of politicians and aims to reflect the "dark factor of personality" suggested by Moshagen et al. (2018). We have described the development and psychometric characteristics of the scale in

detail elsewhere (Maier et al., 2022). The scale includes the following items (in parentheses: represented facets of aversive personality): "There have been times when I was willing to suffer some small harm so that I could punish someone else who deserved it" (spitefulness); "It's wise to keep track of information that you can use against people later" (Machiavellianism); "There are things you should hide from other people to preserve your reputation" (Machiavellianism); "I insist on getting the respect I deserve" (narcissism); "I want my rivals to fail" (narcissism); "People who mess with me always regret it" (psychopathy). Consistent with the concept of the dark core of personality (Moshagen et al., 2018, 2020), the number of traits represented by PEAPS goes beyond the dimensions proposed by the dark triad (i.e., narcissism, psychopathy, and Machiavellianism; see Jones & Paulhus, 2014). All items were measured on a five-point Likert scale from 1 (*do not agree at all*) to 5 (*fully agree*), with only the endpoints of the scale verbalized. Reliability of the scale is Cronbach's α = 0.67.

Obviously, asking politicians about their (aversive) personality traits is not without challenges. First, social desirability might cause candidates scoring high on dark personality traits to not participate in our survey or not reveal their "true" personality. Although candidate studies that have surveyed basic personality traits using the five-factor/big five framework or the HEXACO inventory suggest that politicians attribute more socially desirable characteristics to themselves, the observed bias does not appear to be excessively strong, at least not stronger than in citizen samples (Schumacher & Zettler, 2019). Indeed, it is unclear what qualities the politicians themselves consider desirable; they may consider high self-esteem, tactical skill, and a certain ruthlessness to be prerequisites for real success in the political arena (Schumacher & Zettler, 2019). In line with this consideration, some studies report that politicians score lower than ordinary citizens on some personality traits intuitively rated as socially desirable (e.g., Best, 2011). Second, it is difficult to motivate politicians to answer questions about socially less tolerated characteristics; the risk that they will stop answering the questionnaire is high, particularly if the number of questions is very high (as this is often the case for personality measures). However, by using the presented short scale, we try to minimize these problems.

3.2.2. Independent Variables

To assess the social profile of candidates, we use gender (0 = male, 1 = female) and age, both stemming from data of the state returning officer. Political profile is measured first by whether a candidate was an incumbent, i.e., whether a candidate was a member of parliament before the election (0 = no, 1 = yes), and second by electoral success. This information was also taken from the state returning officer. Furthermore, ideology is measured by the candidates' self-reported left–right position (11-point scale from 1 *left* to 11 *right*). Ideological extremism is measured on a six-point scale from 0 *moderate* to 5 *extreme*, obtained by folding the left–right variable on itself.

3.2.3. Controls

Several meta-studies have investigated the correlations between basic personality traits, measured via the big five framework or the HEXACO inventory, and aversive personality traits, both at the citizen level (e.g., Moshagen et al., 2018; Muris et al., 2017; O'Boyle et al., 2015; Schreiber & Marcus, 2020) but also among politicians (Nai, 2019a, 2019b, 2022; Nai & Martínez i Coma, 2019; Nai et al., 2019; Nai & Toros, 2020; Simonton, 1988). The most consistent patterns exist for agreeableness and, to a somewhat lesser degree, conscientiousness (for both negative correlations with dark personality). We use the 24-item brief HEXACO inventory (De Vries, 2013; Schumacher & Zettler, 2019) to assess basic personality traits. Please note that two items to measure the honesty-humility trait were slightly modified to reduce the risk that politicians stop answering the questionnaire (see Table B1 of Appendix B in the Supplementary File). Furthermore, we have omitted the item "*Ich bin selten aufgeregt*" ("I am seldom cheerful") from the index for extraversion because, contrary to expectations, it correlated negatively with the scale. The reliability values for the HEXACO traits are quite low (e.g., Bakker & Lelkes, 2018; honesty–humility: α = 0.40; emotionality: α = 0.40; extraversion: α = 0.61; agreeableness vs. anger: α = 0.41; conscientiousness: α = 0.50; openness for experience: α = 0.46). One factor that might explain this is that short scales use only a few items for each trait, which in turn measure only a small subset of the characteristics of their subdimensions.

Furthermore, we control for the mode of participation (online vs. paper-and-pencil) and the different elections.

4. Results

Figure 1 shows the distribution of the candidates' self-reported aversive personalities. The distribution almost follows a normal distribution; however, the distribution is slightly skewed to the left, indicating that the average candidate scores somewhat below the mean of the scale. In fact, the mean is M = 2.50 (SD = 0.70), reflecting a moderate level of "dark" personality among politicians. Compared to the study by Nai (2022), who reported a mean score for top politicians worldwide, we find somewhat lower levels of aversive personality. However, it is unlikely that general social desirability is at play here. Compared to the findings of Bader et al. (2021), who analyzed a German student sample using the D-16, the D-35, and the D-70 scales representing the dark core of personality, the average level of aversive personality reported by politicians is about 0.5–0.7 scale points *higher*.

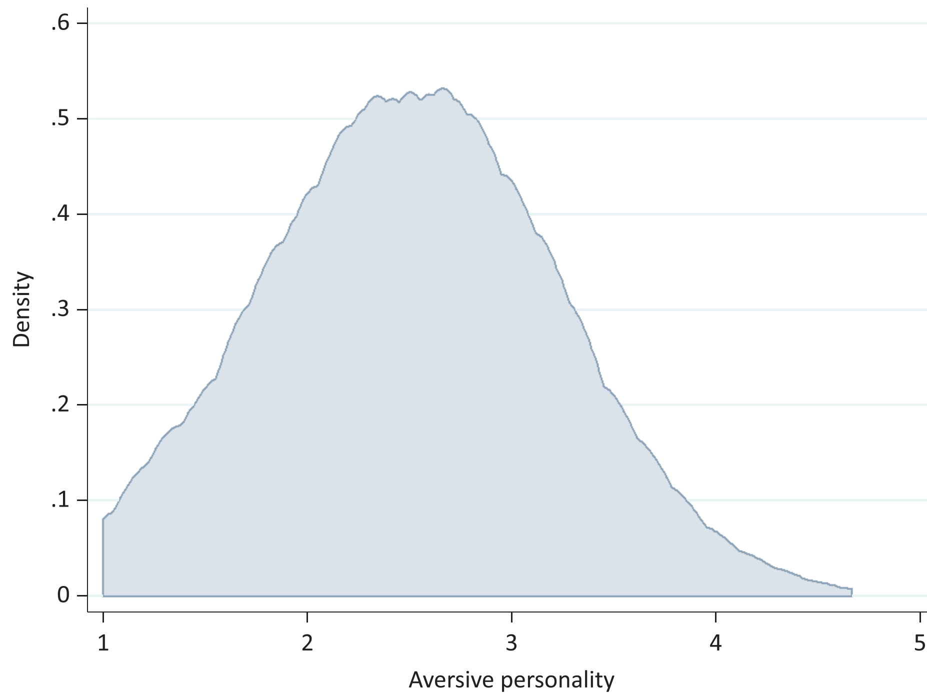

Figure 1. Kernel density plot of self-reported aversive personality traits. Notes: N = 1,441 candidates; Kernel Epanechnikov bandwidth = 0.20.

Can we explain the variation in aversive personality between candidates with their social and political profiles? Model 1 in Table 1 shows that the variables included in our regression model explain only 4.9% of the variation in "dark" personality. Age (b = −0.007, $p < 0.001$), ideology (b = 0.070, $p < 0.001$), and extremism (b = 0.056, $p < 0.001$) are the most powerful predictors of aversive personality traits. Female (b = −0.087, $p < 0.05$)

Table 1. Prediction of self-reported dark personality traits of politicians.

	Model 1		Model 2	
	b	*SE*	*b*	*SE*
Adjusted R^2	0.049		0.270	
Gender	−0.087*	(0.040)	−0.041	(0.038)
Age	−0.007***	(0.001)	−0.005***	(0.001)
Incumbent	0.057	(0.085)	−0.011	(0.076)
Electoral success	−0.127*	(0.060)	−0.025	(0.054)
Ideology	0.070***	(0.012)	0.047***	(0.010)
Extremism	0.056***	(0.016)	0.034*	(0.014)
Honesty–humility			−0.388***	(0.030)
Emotionality			−0.010	(0.025)
Extraversion			−0.049	(0.026)
Agreeableness vs. anger			−0.314***	(0.028)
Conscientiousness			0.023	(0.026)
Openness to experience			0.035	(0.027)
Participation: Paper and pencil			−0.062	(0.042)
Rhineland-Palatinate			0.043	(0.046)
Saxony-Anhalt			0.004	(0.061)
Berlin			0.046	(0.047)
Mecklenburg-Western Pomerania			0.082	(0.069)
Saarland			0.043	(0.070)
Constant	2.549***	(0.115)	5.200***	(0.232)
N	1,321		1,320	

Notes: Displayed are unstandardized coefficients of an OLS regression; SE stands for standard error; significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

and successful (i.e., elected) candidates (b = −0.127, p = 0.05) exhibit significantly lower levels of aversive personality traits. Finally, incumbency (no support for H3) is not significantly correlated with aversive personality.

Not surprisingly, the adjusted R^2 increases sharply when basic personality traits are added, suggesting that aversive personality traits are strongly related to more general personality traits measured via the HEXACO inventory (M2 in Table 1). Bivariate analyses show that the self-reported aversive personality traits are meaningfully correlated with some traits of the HEXACO inventory. We find significant negative correlations with honesty–humility ($r(1,440) = -0.42$, $p < 0.001$), agreeableness vs. anger ($r(1,441) = -0.38$; $p < 0.001$), extraversion ($r(1,441) = -0.12$, $p < 0.001$), and conscientiousness ($r(1,440) = -0.10$, $p < 0.001$). In contrast, the PEAPS scale is uncorrelated with emotionality ($r(1,440) = 0.04$, $p > 0.05$) and openness for experience ($r(1,441) = -0.04$, $p > 0.05$).

More relevant for our research question, however, age (supporting H2), ideology (supporting H5), and extremism (supporting H6) still significantly predict aversive personality after controlling for basic personality traits, the mode of participation, and the different elections (see Model 2 in Table 1). The likelihood of reporting an aversive personality significantly decreases with age. The predicted difference between a candidate aged 20 and 80 is about 0.30 scale points (Figure 2). In addition, the likelihood of aversive personality traits significantly increases the more candidates describe themselves as ideologically right-leaning. The predicted difference between a candidate from the far-left and the far-right is about half a scale point (Figure 3). Furthermore, the likelihood of an aversive personality increases with an extreme ideological position. The predicted difference between a moderate and an extreme candidate is somewhat lower than one-fifth of a scale point (Figure 4). There is no interaction between ideology and extremism, suggesting that it is not right-wing extremism that specifically correlates with dark personality (see Table D2 of Appendix D in the Supplementary File). However, gender (no support for H1) and electoral success (no support for H4) become insignificant after controlling for the HEXACO traits. Our model does not suffer from multicollinearity (see Table D1 of Appendix D in the Supplementary File).

5. Summary and Conclusion

An increasing number of studies examine politicians' personalities based on the assumption that personality, among other factors, can make an important contribution to understanding the behavior of politicians. Socially aversive—but not pathological—personality traits are held responsible for behavior that violates generally accepted ethical, moral, and social norms and thus can threaten democracy. We contribute to this emerging line of research by analyzing *self-reported* aversive (or "dark") personality traits of candidates running for German state parliaments.

Our results suggest that candidates, on average, have moderate levels of aversive personality traits. The level found in our data is somewhat lower than that reported for top politicians (Nai, 2022) but clearly higher than for citizen samples. This suggests that social desirability is not a severe problem in our case, which is consistent with other research measuring self-reported personality traits in politicians (Best, 2011; Schumacher & Zettler, 2019). This makes sense intuitively as it is unclear what qualities politicians themselves consider desirable; they

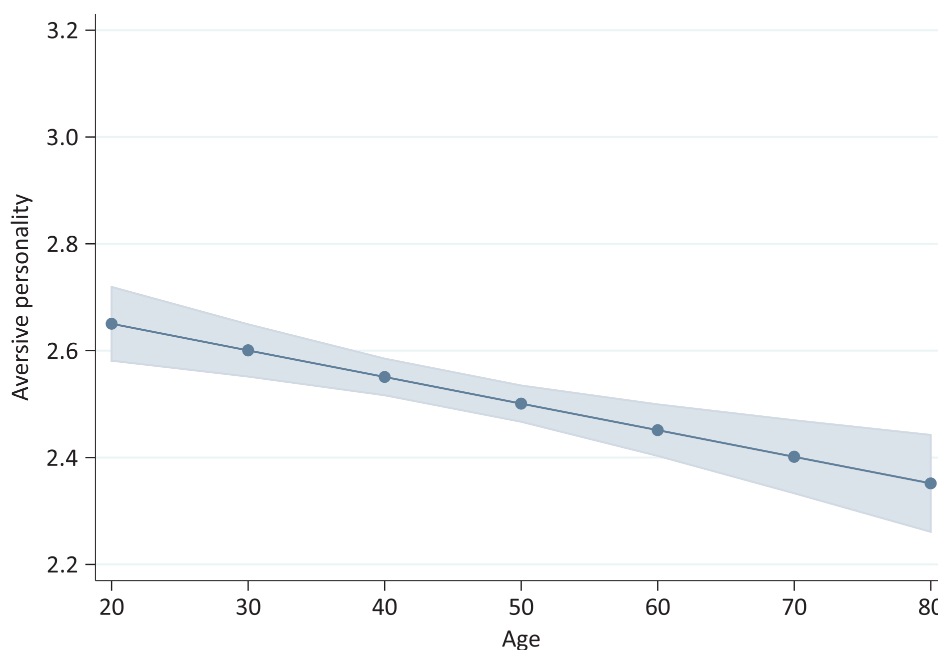

Figure 2. Predicted margins of age on aversive personality with 95% confidence intervals based on M2.

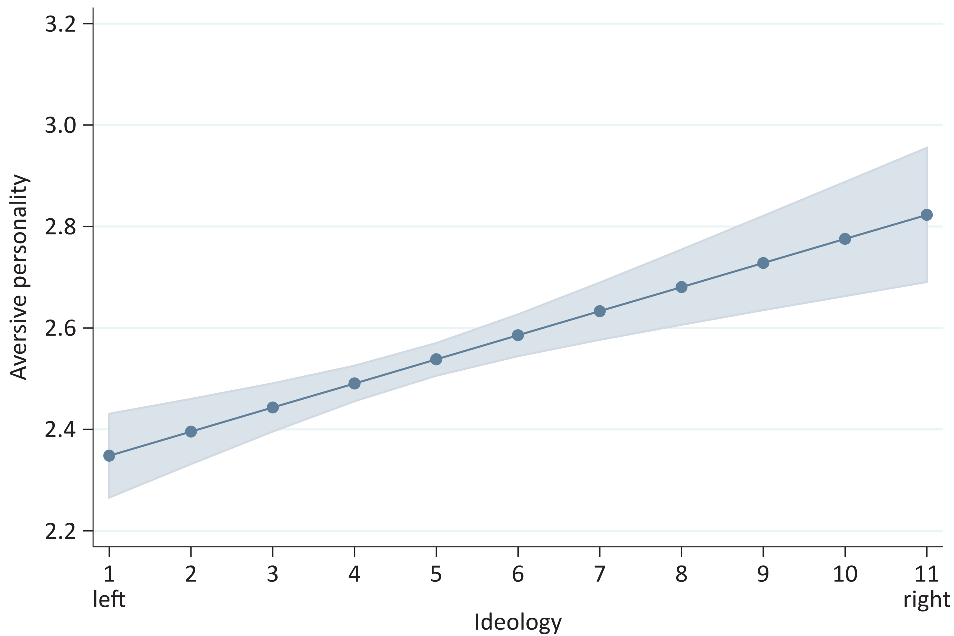

Figure 3. Predicted margins of ideology on aversive personality with 95% confidence intervals based on M2.

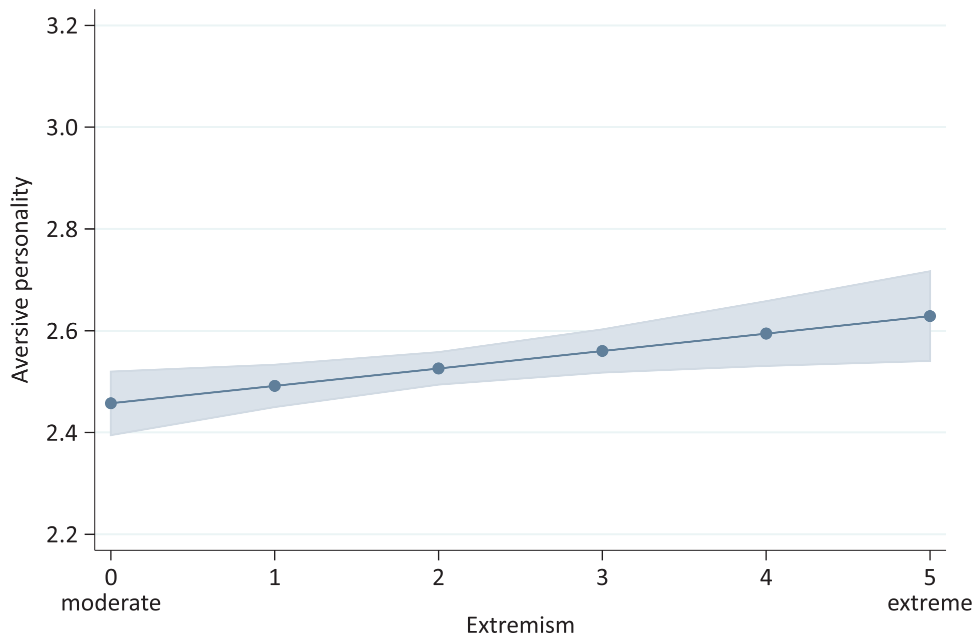

Figure 4. Predicted margins of ideological extremism on aversive personality with 95% confidence intervals based on M2.

may consider high self-esteem, tactical skill, and a certain ruthlessness to be prerequisites for real success in the political arena (Schumacher & Zettler, 2019). Although we have no direct empirical evidence that our data suffer from social desirability, the extent of dark personality among politicians that we report could nevertheless be a conservative estimate. This is supported on the one hand by the fact that aversive personality is associated with extreme ideology and—even more strongly—with self-positioning as ideologically right, and on the other hand by the fact that candidates from CDU and AfD, i.e., parties ideologically on the right (Dippel & Burger, 2022), participate less frequently in our surveys (see Table C1 of Appendix C in the Supplementary File).

In addition, we find that the candidates' social and political profiles predict their self-reported level of "dark" personality. Aversive personality traits are significantly more likely for younger candidates, those who classify themselves as conservative, and the ideologically extreme. In contrast, we do not find differences between male and female candidates, incumbents and challengers, and those who have won a seat in the

parliament and those who have not. The results suggest that candidates with a high level of aversive personality are not more successful per se than less "dark" politicians. Our results further indicate that candidates with certain ideological positions (i.e., right-wing candidates, politically extreme candidates) are more likely to show aversive personality traits. At a time when polarization (i.e., the strengthening of the political fringes) is advancing, it is increasingly likely that "dark" candidates will enter the political arena. They might not be successful on election day, but they can, of course, harm the political process if their communication which is more likely to be negative, uncivil, populist, or based on fake information, attracts the media (e.g., Maier & Nai, 2020; Maurer et al., 2022), and a particular segment of voters (Nai & Maier, 2021b). Moreover, our results suggest that age predicts aversive personality traits. The average age in the federal parliament has decreased in the current legislative term (Feldkamp, 2022; information for state parliaments is unfortunately unavailable); hence, it is expected that the proportion of "dark" politicians has thus increased. Finally, our results show that aversive personality traits are meaningfully correlated with basic personality traits. Higher levels of "dark" personality go hand in hand with low agreeableness (vs. anger), low honesty-humility, and low extraversion.

Our approach comes with some limitations that open potential for future studies. First, our study focuses on German politicians. Comparative research is warranted to see whether there are differences concerning the distribution and the correlates of aversive personality traits in other countries. Second, our study is based on candidates running for German state parliaments. It would be interesting to compare our results with members of the national parliament. Third, our data only provide information on a limited number of candidate characteristics. Hence, future studies should include more variables to get a better picture of what predicts the level of aversive personality traits among political elites. Fourth, although we asked politicians to complete the questionnaire themselves, we have no way of verifying whether they did so. It could be that, in some cases, the survey was conducted by their staff. Fifth, the reliability of the scales we use is not particularly high. This is less true for PEAPS, which has sufficient reliability, than the HEXACO traits. One factor that might explain this is that short scales use only a few items for each trait, which in turn measure only a small subset of the characteristics of their subdimensions. Sixth, the focus of our study was limited to the distribution and the correlates of aversive personality traits. It would be very interesting to analyze the consequences of a "dark" personality, for instance, regarding candidates' campaign communication or, once in office, their policy accomplishments. Finally, we can only make very few comparisons between political elites and the population. However, with regard to questions of descriptive representation, representative surveys on the distribution and determinants of "dark" personality in the electorate would be very important. Our article and the short scale used to measure the core of aversive personality traits set the stage for such research.

Acknowledgments

Jürgen Maier and Corinna Oschatz acknowledge financial support from the German Research Foundation (DFG; Grant No. 441574527). The GESIS ethics committee approved the study on 27 November 2020 (Reference No. 2020–6). The authors would like to thank all the candidates who took the time to participate in our surveys despite their busy schedules.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*(2), 150–166. https://doi.org/10.1177/1088868306294907

Bader, M., Horsten, L. K., Hilbig, B. E., Zettler, I., & Moshagen, M. (2021). Measuring the dark core of personality in Germany: Psychometric properties, measurement invariance, predictive validity, and self-other agreement. *Journal of Personality Assessment*, *104*(5), 660–673. https://doi.org/10.1080/00223891.2021.1984931

Bakker, B. N., & Lelkes, Y. (2018). Selling ourselves short? How abbreviated measures of personality change the way we think about personality and politics. *Journal of Politics*, *80*(4), 1311–1325. https://doi.org/10.1086/698928

Bannon, B. (2020, October 10). Trump's personality is as much a problem as his performance. *The Hill*. https://thehill.com/opinion/campaign/511368-trumps-personality-is-as-much-a-problem-as-his-performance

Best, H. (2011). Does personality matter in politics? Personality factors as determinants of parliamentary recruitment and policy preferences. *Comparative Sociology*, *10*(6), 928–948. https://doi.org/10.1163/156913311X607638

Bildt, C. (2018, October 30). The end of an era in Europe. *The Washington Post*. https://www.washingtonpost.com/news/global-opinions/wp/2018/10/30/the-end-of-an-era-in-europe

Blais, J., & Pruysers, S. (2017). The power of the dark side: Personality, the dark triad, and political ambition.

Personality and Individual Differences, *113*, 167–172. https://doi.org/10.1016/j.paid.2017.03.029

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science*, *24*(11), 2201–2209. https://doi.org/10.1177/0956797613490749

Caprara, G. V., & Vecchione, M. (2017). *Personalizing politics and realizing democracy*. Oxford University Press.

Carstensen, L. L., Fung, H. H., & Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motivation and Emotion*, *27*(2), 103–123. https://doi.org/10.1023/A:1024569803230

Chabrol, H., Bronchain, J., Morgades Bamba, C. I., & Raynal, P. (2020). The dark tetrad and radicalization: Personality profiles in young women. *Behavioral Sciences of Terrorism and Political Aggression*, *12*(2), 157–168. https://doi.org/10.1080/19434472.2019.1646301

Chabrol, H., Van Leeuwen, N., Rodgers, R., & Séjourné, N. (2009). Contributions of psychopathic, narcissistic, Machiavellian, and sadistic personality traits to juvenile delinquency. *Personality and Individual Differences*, *47*(7), 734–739. https://doi.org/10.1016/j.paid.2009.06.020

Chopik, W. J., & Kitayama, S. (2018). Personality change across the life span: Insights from a cross-cultural, longitudinal study. *Journal of Personality*, *86*(3), 508–521. https://doi.org/10.1111/jopy.12332

De Vries, R. E. (2013). The 24-item brief HEXACO inventory (BHI). *Journal of Research in Personality*, *47*(6), 871–880. https://doi.org/10.1016/j.jrp.2013.09.003

Dietrich, B. J., Lasley, S., Mondak, J. J., Remmel, M. L., & Turner, J. (2012). Personality and legislative politics: The big five trait dimensions among US state legislators. *Political Psychology*, *33*(2), 195–210. https://doi.org/10.1111/j.1467-9221.2012.00870.x

Dippel, A. S., & Burger, A. M. (2022). Links oder rechts? Die ideologische Selbstverortung von Wählerinnen und ihre Wahrnehmung von Parteien in Deutschland [Left or right? The ideological self-placement of voters and their perception of parties in Germany]. *easy_social_sciences*, *67*, 19–29. https://doi.org/10.15464/easy.2022.04

Duckitt, J. (2006). Differential effects of right wing authoritarianism and social dominance orientation on outgroup attitudes and their mediation by threat from and competitiveness to outgroups. *Personality and Social Psychology Bulletin*, *32*(5), 684–696. https://doi.org/10.1177/0146167205284282

Feldkamp, M. F. (2022). Deutscher Bundestag 2002 bis 2021/22: Parlaments- und Wahlstatistik für die 15. bis beginnende 20 Wahlperiode [Deutscher Bundestag 2002–2021/22: Parliamentary and electoral statistics for the 15th to the beginning 20 electoral period]. *Zeitschrift für Parlamentsfragen*, *53*(2), 243–260. https://doi.org./10.5771/0340-1758-2022-2-243

Hanania, R. (2017). The personalities of politicians: A big five survey of American legislators. *Personality and Individual Differences*, *108*, 164–167. https://doi.org/10.1016/j.paid.2016.12.020

Hartung, J., Bader, M., Moshagen, M., & Wilhelm, O. (2022). Age and gender differences in socially aversive ("dark") personality traits. *European Journal of Personality*, *36*(1), 3–23. https://doi.org/10.1177/0890207020988435

Joly, J., Soroka, S., & Loewen, P. (2019). Nice guys finish last: Personality and political success. *Acta Politica*, *54*(4), 667–683. https://doi.org/10.1057/s41269-018-0095-z

Jones, D. N. (2013). Psychopathy and Machiavellianism predict differences in racially motivated attitudes and their affiliations. *Journal of Applied Social Psychology*, *43*(S2), E367–E378. https://doi.org/10.1111/jasp. 12035

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3): A brief measure of dark personality traits. *Assessment*, *21*(1), 28–41. https://doi.org/10.1177/1073191113514105

Kajonius, P. J., Persson, B. N., & Jonason, P. K. (2015). Hedonism, achievement, and power: Universal values that characterize the dark triad. *Personality and Individual Differences*, *77*, 173–178. https://doi.org/10.1016/j.paid.2014.12.055

Klimstra, T. A., Jeronimus, B. F., Sijtsema, J. J., & Denissen, J. J. (2020). The unfolding dark side: Age trends in dark personality features. *Journal of Research in Personality*, *85*, Article 103915. https://doi.org/10.1016/j.jrp. 2020.103915

Landay, K., Harms, P. D., & Credé, M. (2019). Shall we serve the dark lords? A meta-analytic review of psychopathy and leadership. *Journal of Applied Psychology*, *104*(1), 183–196. https://doi.org/10.1037/apl0000357

Leiner, D. J. (2019). Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. *Survey Research Methods*, *13*(3), 229–248. https://doi.org/10.18148/SRM/2019.V13I3.7403

Lilienfeld, S. O., Waldman, I. D., Landfield, K., Watts, A. L., Rubenzer, S., & Faschingbauer, T. R. (2012). Fearless dominance and the US presidency: Implications of psychopathic personality traits for successful and unsuccessful political leadership. *Journal of Personality and Social Psychology*, *103*(3), 489–505. https://doi.apa.org/doi/10.1037/a0029392

Lyons, M. (2019). *The dark triad of personality: Narcissism, Machiavellianism, and psychopathy in everyday life*. Academic Press.

Maier, J., & Nai, A. (2020). Roaring candidates in the spotlight: Campaign negativity, emotions, and media coverage in 107 national elections. *International Journal of Press/Politics*, *25*(4), 576–606. https://doi.org/10.1177/1940161220919093

Maier, J., & Nai, A. (2021). Mapping the drivers of negative campaigning: Insights from a candidate sur-

vey. *International Political Science Review*. Advance online publication. https://doi.org/10.1177/01925 12121994512

Maier, J., Oschatz, C., Stier, S., & Zettler, I. (2022). A short scale to measure self-reported aversive personality traits in political elites. *Journal of Personality Assessment*. Advance online publication. https://doi.org/10.1080/00223891.2022.2130341

Maurer, M., Jost, P., Schaaf, M., Sülflow, M., & Kruschinski, S. (2022). How right-wing populists instrumentalize news media: Deliberate provocations, scandalizing media coverage, and public awareness for the Alternative for Germany (AfD). *International Journal of Press/Politics*. Advance online publication. https://doi.org/10.1177/19401612211072692

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Moshagen, M., Hilbig, B. E., & Zettler, I. (2018). The dark core of personality. *Psychological Review*, *125*(5), 656–688. https://doi.org/10.1037/rev0000111

Moshagen, M., Zettler, I., & Hilbig, B. E. (2020). Measuring the dark core of personality. *Psychological Assessment*, *32*(2), 182–196. https://doi.org/10.1037/pas0000778

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science*, *12*(2), 183–204. https://doi.org/10.1177/1745691 616666070

Nai, A. (2019a). Disagreeable narcissists, extroverted psychopaths, and elections: A new dataset to measure the personality of candidates worldwide. *European Political Science*, *18*(2), 309–334. https://doi.org/10.1057/s41304-018-0187-2

Nai, A. (2019b). The electoral success of angels and demons: Big five, dark triad, and performance at the ballot box. *Journal of Social and Political Psychology*, *7*(2), 830–862. https://doi.org/10.5964/jspp.v7i2.918

Nai, A. (2022). Populist voters like dark politicians. *Personality and Individual Differences*, *187*, Article 111412. https://doi.org/10.1016/j.paid.2021.111412

Nai, A., & Maier, J. (2018). Perceived personality and campaign style of Hillary Clinton and Donald Trump. *Personality and Individual Differences*, *121*, 80–83. https://doi.org/10.1016/j.paid.2017.09.020

Nai, A., & Maier, J. (2020). Dark necessities? Candidates' aversive personality traits and negative campaigning in the 2018 American Midterms. *Electoral Studies*, *68*, Article 102233. https://doi.org/10.1016/j.electstud.2020.102233

Nai, A., & Maier, J. (2021a). Can anyone be objective about Donald Trump? Assessing the personality of political figures. *Journal of Elections, Public Opinion and Parties*, *31*(3), 283–308. https://doi.org/10.1080/17457289.2019.1632318

Nai, A., & Maier, J. (2021b). Is negative campaigning a matter of taste? Political attacks, incivility, and the moderating role of individual differences. *American Politics Research*, *49*(3), 269–281. https://doi.org/10.1177/1532673X20965548

Nai, A., & Martínez i Coma, F. (2019). The personality of populists: Provocateurs, charismatic leaders, or drunken dinner guests? *West European Politics*, *42*(7), 1337–1367. https://doi.org/10.1080/01402382.2019.1599570

Nai, A., Martínez i Coma, F., & Maier, J. (2019). Donald Trump, populism, and the age of extremes: Comparing the personality traits and campaigning styles of Trump and other leaders worldwide. *Presidential Studies Quarterly*, *49*(3), 609–643. https://doi.org/10.1111/psq.12511

Nai, A., & Toros, E. (2020). The peculiar personality of strongmen: Comparing the big five and dark triad traits of autocrats and non-autocrats. *Political Research Exchange*, *2*(1), Article 1707697. https://doi.org/10.1080/2474736X.2019.1707697

O'Boyle, E. H., Forsyth, D. R., Banks, G. C., Story, P. A., & White, C. D. (2015). A meta-analytic test of redundancy and relative importance of the dark triad and five-factor model of personality. *Journal of Personality*, *83*(6), 644–664. https://doi.org/10.1111/jopy.12126

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*(3), 913–938. https://doi.org/10.1017/S0954579 409000492

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*(6), 556–563. https://doi.org/10.1016/S0092-6566(02)00505-6

Pavlović, T., & Wertag, A. (2021). Proviolence as a mediator in the relationship between the dark personality traits and support for extremism. *Personality and Individual Differences*, *168*, Article 110374. https://doi.org/10.1016/j.paid.2020.110374

Ramey, A. J., Klingler, J. D., & Hollibaugh, G. E. (2019). Measuring elite personality using speech. *Political Science Research and Methods*, *7*(1), 163–184. https://doi.org/10.1017/psrm.2016.12

Rauthmann, J. F. (2012). The dark triad and interpersonal perception: Similarities and differences in the social consequences of narcissism, Machiavellianism, and psychopathy. *Social Psychological and Personality Science*, *3*(4), 487–496. https://doi.org/10.1177/1948550611427608

Rice, M. G., Remmel, M. L., & Mondak, J. J. (2021). Personality on the hill: Expert evaluations of US senators' psychological traits. *Political Research*

Quarterly, *74*(3), 674–687. https://doi.org/10.1177/1065912920928587

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–39). Erlbaum.

Rubenzer, S. J., & Faschingbauer, T. R. (2004). *Personality, character, and leadership in the White House: Psychologists assess the presidents*. Brassey's.

Rubenzer, S. J., Faschingbauer, T. R., & Ones, D. S. (2000). Assessing the US presidents using the revised NEO personality inventory. *Assessment*, *7*(4), 403–419. https://doi.org/10.1177/107319110000700408

Schmitt, D. P., Long, A. E., McPhearson, A., O'Brien, K., Remmert, B., & Shah, S. H. (2017). Personality and gender differences in global perspective. *International Journal of Psychology*, *52*(S1), 45–56. https://doi.org/10.1002/ijop. 12265

Schreiber, A., & Marcus, B. (2020). The place of the "dark triad" in general models of personality: Some meta-analytic clarification. *Psychological Bulletin*, *146*(11), 1021–1041. https://doi.org/10.1037/bul0000299

Schumacher, G., & Zettler, I. (2019). *House of Cards* or *West Wing*? Self-reported HEXACO traits of Danish politicians. *Personality and Individual Differences*, *141*, 173–181. https://doi.org/10.1016/j.paid.2018.12.028

Scott, C., & Medeiros, M. (2020). Personality and political careers. What personality types are likely to run for office and get elected? *Personality and Individual Differences*, *152*, Article 109600. https://doi.org/10.1016/j.paid.2019.109600

Simonton, D. K. (1988). Presidential style: Personality, biography, and performance. *Journal of Personality and Social Psychology*, *55*(6), 928–936. https://doi.org/10.1037/0022-3514.55.6.928

Van Hiel, A., & Mervielde, I. (2002). Explaining conservative beliefs and political preferences: A comparison of social dominance orientation and authoritarianism. *Journal of Applied Social Psychology*, *32*(5), 965–976. https://doi.org/10.1111/j.1559-1816.2002.tb00250.x

Visser, B. A., Book, A. S., & Volk, A. A. (2017). Is Hillary dishonest and Donald narcissistic? A HEXACO analysis of the presidential candidates' public personas. *Personality and Individual Differences*, *106*, 281–286. https://doi.org/10.1016/j.paid.2016.10.053

Vize, C. E., Collison, K. L., Miller, J. D., & Lynam, D. R. (2020). The "core" of the dark triad: A test of competing hypotheses. *Personality Disorders: Theory, Research, and Treatment*, *11*(2), 91–99. https://doi.org/10.1037/per0000386

Vladimir Putin: Russia's action man president. (2021, April 16). *BBC News*. https://www.bbc.com/news/world-europe-15047823

Watts, A. L., Lilienfeld, S. O., Smith, S. F., Miller, J. D., Campbell, W. K., Waldman, I. D., Rubenzer, S. J., & Faschingbauer, T. J. (2013). The double-edged sword of grandiose narcissism: Implications for successful and unsuccessful leadership among US presidents. *Psychological Science*, *24*(12), 2379–2389. https://doi.org/10.1177/0956797613491970

About the Authors

Jürgen Maier is a professor of political communication at the Department of Political Science, University of Koblenz-Landau. His research focuses on the content and the impact of campaign communication.

Mona Dian is a research assistant at the Department of Political Science, University of Koblenz-Landau. She is interested in political sociology and individual differences in political behavior, especially the role of personality in election campaigns.

Corinna Oschatz is an assistant professor of political communication and journalism at the Amsterdam School of Communication Research, University of Amsterdam. She holds a PhD in communication science from the University of Mainz (2016). Her research focuses on political communication, persuasion, and methods of empirical social science.

cogitatio

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 361–373
https://doi.org/10.17645/pag.v10i4.5723

Article

Do Leader Evaluations (De)Mobilize Voter Turnout? Lessons From Presidential Elections in the United States

Liran Harsgor [1,*] and Neil Nevitte [2]

[1] School of Political Science, University of Haifa, Israel
[2] Department of Political Science, University of Toronto, Canada

* Corresponding author (lharsgor@poli.haifa.ac.il)

Submitted: 30 April 2022 | Accepted: 8 September 2022 | Published: 30 December 2022

Abstract
Do evaluations of presidential candidates in the US affect the level of voter turnout? Voters' affections towards presidential candidates, we contend, can either stimulate or inhibit voter inclinations to turnout. Voters are more inclined to turn out when they have positive feelings towards the candidate with which they identify because they want "their" candidate to win. But citizens may also be more likely to vote when they dislike the candidate of the party with which they do not identify. In that case, voters are motivated to prevent the candidate from being elected. Utilizing the American National Election Studies data for 1968–2020, the analysis finds that the likelihood of voting is affected by (a) the degree to which voters' affections towards the candidate differ from one another (having a clear-cut choice between options) and (b) the nature of the affections (negative or positive) towards both in- and out-party candidates.

Keywords
leader evaluations; mobilization; presidentialization; turnout; US elections

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Leaders matter more now to electoral politics than they once did. This shift towards leader-centered electoral politics signifies the growing power of leaders vis-à-vis their political parties (Elgie & Passarelli, 2018; Poguntke & Webb, 2005; Samuels & Shugart, 2010). This has been accompanied by the rise in candidate-centered campaigns and a greater focus on leader evaluations and personalities (Balmas et al., 2014). The significant comparative point is that these dynamics, which are characterized as the presidentialization of electoral politics, operate across different electoral systems and apply to parliamentary and presidential style regimes alike (Mughan, 2000; Norton, 2003; Poguntke & Webb, 2005; Rahat & Kenig, 2018). One conjecture is that the shifts to the presidentialization of politics are attributable, among other things, to the joint effects of the erosion of the structural and ideological underpinnings of traditional political parties and the changing character of mass communications (Poguntke & Webb, 2005). The following investigation focuses on the American context, drawing on more than 50 years' worth of data from the American National Election Studies (ANES), and examines whether and how leader evaluations may have mobilized or demobilized voter turnout in that setting.

The article makes two main arguments: The first concerns the degree to which voters differentiate between candidates. Studies have demonstrated that voters who have clear ideological or affective evaluations of their electoral parties participate more than indifferent voters. But voters also care about political leaders when

casting their ballot. Therefore, we suggest that citizens with distinctive leader evaluations are more likely to vote than those who hold similar evaluations for the two candidates. The second argument considers the relationship between leader evaluations, partisanship, and voter turnout. Voters' affection towards party leaders, we contend, can serve as both a stimulating and inhibiting factor when it comes to turnout. The more nuanced conjecture is that the direction of the effect depends on party affiliation. Thus, we hypothesize about the differences between the effects of in-party leaders' evaluations and out-party leaders' evaluations on voters' turnout. Leaders can increase voting when voters have positive feelings towards the leader of the party they identify with because voters would want "their" candidate to win. Citizens may also be more likely to vote when they dislike the candidate of the out-party, as they want to prevent him/her from being elected. We attempt to evaluate the effect of affections towards both the in- and out-party candidates and their interactive effect.

We begin the article by describing the common predictors for voter turnout at the macro and the micro levels, then moving to portray the theoretical underpinnings of the assumed effects of leader evaluations on turnout. After introducing the data and variable measures, the analysis proceeds in three stages. The first examines the effect on the turnout of having a clear-cut choice between candidates. The second deals with evaluations of in- and out-party leaders as well as among independents and their effects on turnout. The third part delves into the more nuanced picture of interacted effects of in- and out-party leader evaluations. Our findings show that candidate evaluations clearly do matter for levels of turnout. Moreover, the effects vary according to partisanship and the type of affection the voter exhibits towards both her in- and out-party candidates. The analysis demonstrates that leader evaluations play a key role in mobilizing voters to the polls. It shows that citizens with a clear-cut choice between candidates tend to vote more than others and that people are mobilized to vote when they hold positive evaluations of their in-party candidate. Significantly, negative feelings for the out-party candidate mobilize voters only when they feel positively about their in-party candidate. The presidentialization hypothesis encourages the expectation that these effects may have become more magnified with the passage of time. The data, however, do not provide clear support for that contention. The article concludes by considering some of the broader implications of the findings.

2. Leader Evaluations and Turnout: Why and How?

2.1. Explanations for Voter Turnout

Voter turnout stood at 63% in the 1960 US presidential elections. That gradually dropped to 50% by 1996 and then experienced a modest uptick in 2020. Explanations for variations in voter participation rates typically focus either on aggregate system-level characteristics or individual-level micro-foundations (Smets & van Ham, 2013). The pioneering investigations of Powell (1986) and Jackman (1987) both made the strong case that institutions decisively shape voter turnout rates. Jackman (1987) identified five features he thought to be particularly important: multipartyism, proportional representation (PR) electoral rules, unicameralism, electoral competitiveness, and compulsory voting. These seemed intuitively reasonable. PR systems appear "fairer," and multipartyism—the product of PR electoral rules—plausibly increases voter participation because citizens are presented with more choices. Few of these initial propositions, however, remain intact after subsequent scrutiny. Multiple studies show that the impact of PR electoral rules and multiparty arrangements on turnout is mixed at best (Blais, 2006; Geys, 2006). Rather, the presence of more parties makes the choices facing voters more complicated and outcomes harder to predict. The impact of unicameralism on turnout is also mixed. Jackman and Miller (1995) report positive effects, but then others (Blais & Carty, 1990; Radcliff & Davis, 2000) find no such effects whatsoever. Nor is there clear evidence that turnout is higher in federated systems (Stockemer, 2016). Compulsory voting rules do boost turnout but only in "old" democracies and when accompanied by sanctions (Norris, 1999). As rational choice theories would predict, turnout is indeed higher in "small" countries, when electoral districts are "small," and when elections are competitive (Franklin, 2004).

Macro-level considerations may well help to explain cross-national variations in voter turnout, but they are ill-equipped to account for within-country variations. In the case of the US, for example, federal structures have not changed, and the electoral rules and registration requirements are stable, as is the two-party system. And by no reasonable measure has the US qualified as "small." If macro-level considerations do not plausibly account for variations within country turnout, then which micro-considerations are likely important? There are several socioeconomic factors that have been amply demonstrated to be consistent predictors of individual-level variations in voter participation rates. And none is more important than education (Brady et al., 1995). Citizens with higher levels of formal participation vote more than their lesser-educated counterparts. Education is consistently related to efficacy and interest in politics. Interest, in turn, supplies the motivation to become informed, and it lowers the costs of participation. It comes as no surprise to find that education is also related, typically, to income levels, and wealthier citizens participate more than poorer ones (Verba et al., 1978). Also well documented is a persistent gender gap. That gap may be closing, but the prevailing finding is that men participate more than women, and women are substantially less likely than men to identify with a political party (Inglehart et al., 2003). Party identification, in turn, has a powerful impact both on vote choice and turnout. It provides

an affective template that helps citizens navigate their political worlds, and it provides partisans with information shortcuts about likes and dislikes (Bartels, 2000). Age also matters. Those at the front and back ends of the life cycle participate least; voting is more of a middle age sport (Glenn & Grimes, 1968). History counts, too: Voting in prior elections is a strong predictor of contemporary voter turnout (Fowler, 2006). The clear implication flowing from these collective results is that estimations of whether and how leader evaluations are related to voter turnout need to take into account these microfoundations as well as citizens' partisan predilections.

2.2. Leader Evaluations, Partisanship, and Turnout: Theoretical Expectations

Much of the empirical evidence for leader effects on voting behavior has focused on vote choice (Aarts et al., 2011; Barisione, 2009; Garzia, 2012; Garzia & da Silva, 2021; Gattermann & de Vreese, 2022). These studies have demonstrated that voters consider leader-related calculations when making their electoral choices. However, if voters think of leaders and how they feel about them when casting their ballots, then these leader evaluations should also affect citizens' motivation to go to vote in the first place. Thus, leaders may play an important role in mobilizing—or demobilizing—citizens by motivating them to vote for positively evaluated leaders or against negatively evaluated leaders. Yet, despite the increasing role of leaders in electoral campaigns and voters' considerations, only little attention has been paid to what impact leader evaluations might have on voter turnout. Some recent evidence shows that leader evaluations may have had an increased impact on voter turnout in some European settings (da Silva, 2018; da Silva & Costa, 2019; da Silva et al., 2021). These studies indicate the increasing effect of citizens' evaluations of party leaders on the likelihood of voting, side by side with the declining effect of parties' evaluations. What is called for is a deeper investigation of the effect of leader evaluations in other contexts and across a much longer time span.

The idea that evaluative distances between voters matter derives from spatial theories that focus on ideological distance or issue positions (Lefkofridi et al., 2014; Simas & Ozer, 2021). Those same considerations have also been extended to candidate thermometers (Adams et al., 2006; Brody & Grofman, 1982; da Silva et al., 2021), although the findings from those investigations are inconclusive. There are several ways by which leader evaluations could be related to turnout. We begin by arguing that the way voters feel towards one leader vs. other leaders may indicate the degree to which voters face a clear-cut choice between candidates. If voters like one leader but dislike the other, then the decision of whom to vote for is relatively straightforward. The expectation is that voting is more likely because the choice is easy as the predicted benefit for the voter from one choice over the other is clear (Downs, 1957). In cases where the voters have similar feelings towards the candidates, the scenarios are more complicated. If voters like or dislike both candidates, then electoral outcomes may be seen to be less consequential or important. For these *indifferent* citizens, there is less motivation to vote than for those voters who have sharply different evaluations of the candidates (Adams et al., 2006; Brody & Page, 1973). It could be, though, that disliking both candidates may lead citizens to abstain to a greater degree, as these *alienated* citizens lack the most basic motive to show at the polls (Brody & Grofman, 1982; Weisberg & Grofman, 1981). Our first hypothesis, therefore, concerns the impact of having a clear-cut choice on voter turnout. The empirical analysis examines this effect and compares different groups of voters: the indifferent, the alienated, and those whose evaluations are neither positive nor negative.

> H1: Citizens with distinctive evaluations of the two candidates will be more likely to vote than those who hold similar evaluations of the candidates.

Another question to ask is whether and how the effect of leader evaluations might depend on party affiliation. "Political leaders enter and exit the public stage, but the parties and their symbols, platforms, and group associations provide a long-term anchor to the political system" (Lavine et al., 2012, p. 2). There are several approaches to the concept of party identification, with some scholars seeing it as a very stable trait (Converse, 1969) and others as a running tally of evaluations (Fiorina, 1981). Our main point concerns the way the effect of leader evaluations on turnout depends on the party the voter identifies with. We follow the idea of partisanship as a social identity (Tajfel et al., 1971), which is accompanied by a sense of belonging to the in-group. This, in turn, leads to both in-group and out-group biases as well as different affections towards in- and out-groups' members (Huddy & Bankert, 2017; Iyengar et al., 2012). The analysis of the effect of leader evaluations, we contend, must take party identification into account because positive and negative feelings towards in- and out-party party leaders might have asymmetrical effects on voter turnout. In the past, positive feelings about political parties seem to have had a greater mobilizing effect on participation. But the contemporary evidence seems to be that feelings towards the out-party are more likely to increase voter participation (Iyengar & Krupenkin, 2018). That same general finding also extends to the case of leader evaluations (da Silva et al., 2021). We hypothesize that both positive and negative feelings can mobilize citizens and that these effects depend on party identification. Positive feelings towards the in-party candidate may stimulate voters, while negative feelings towards their in-party candidate can have an inhibiting effect on voting. This is because voters who dislike their in-party candidate may feel cross-pressured due to their party vs. candidate preferences.

Similarly, negative feelings towards the out-party candidate can motivate citizens to vote to stop their disliked candidate from being elected, while liking the candidate of the out-party might lead these cross-pressured citizens to abstain. This is especially true in an era of polarization, in which voters may be less willing to consider voting for the candidate of the other party (Bankert, 2021).

> H2a: The more citizens like the in-party candidate *ceteris paribus*, the greater the likelihood that they vote.

> H2b: The more citizens dislike the out-party candidate *ceteris paribus*, the greater the likelihood that they vote.

What about those who do not identify with a party? Should their leaders' evaluations affect turnout, and if so, how? It is reasonable to conjecture that leader effects might be stronger among independents, those who do not identify with a political party. Certainly, da Silva (2018) finds that leader evaluations have a stronger effect on turnout among independents in a variety of settings. Positive feelings among independents can boost turnout, but so can negative feelings (Bankert, 2022). A variety of voting studies that report on the impact of leader effects and partisanship on turnout also report mixed results. Some report stronger leader effects among non-partisans, while others find stronger leader effects among party identifiers (Gidengil, 2011; Lobo, 2014; Mughan, 2009). Institutional settings capture long-term factors, and it might well be that the mixed findings reflect the impact of short-term leader effects factors. Different leaders, after all, compete across different elections. In the absence of clear guidance from prior results, we proceed cautiously and regard our approach as exploratory.

Lastly, we take a more nuanced approach to examine not only the effect of in-party and out-party evaluations separately but also how they interact with each other. While H2a and H2b speculate about the effect of leader evaluations when controlling for the evaluation of the other candidate, there might be an interacted effect here. The idea is that negative feelings towards the out-party can stimulate turnout, but this effect—or its size—may depend on how people feel for their own candidate. A voter who *dislikes* her in-party candidate, for example, might not be as motivated to vote by having negative feelings towards the out-party candidate as much as a voter who does like her in-party candidate. Thus, negative feelings towards the out-party candidate can motivate citizens to vote, but this might depend on the degree to which these citizens feel comfortable with their in-party candidate.

> H3: The effect of the out-party's leader evaluation on turnout depends on the in-party's leader evaluation, i.e., the more citizens dislike their out-party candidate, the more likely they are to vote, and this effect will get bigger the more they like their in-party candidate.

Exploring these conjectures in the American setting has a number of conceptual and practical advantages. First, the US qualifies as a stable two-party system. This means that voters have faced consistent partisan choices over a long duration. Together, these two attributes encourage relatively stable patterns of partisanship, and if there is one thing about which students of elections agree, it is that partisanship matters. Second, American presidential elections are candidate-centered, and so it is reasonable to suppose that candidate evaluations are likely to have a greater impact in that electoral context. Third, as a practical matter, the ANES have consistently used the very same key measures of such variables as candidate evaluations, party identification, and voter turnout over the duration of these studies. This means that findings based on such indicators are likely to be robust against variations in instrumentation effects.

3. Empirical Analysis: What are the Different Impacts of Leader Evaluations on Turnout?

3.1. Data and Methods

The analysis relies on data from ANES Time Series Cumulative Data File between 1968 and 2020. Additional information can be found online (https://electionstudies.org/data-center/anes-time-series-cumulative-data-file). The empirical investigation has two main independent variables. The first is the presidential candidates' thermometer score. These measures reflect the degree to which voters have warm vs. cold feelings towards each presidential candidate. The question wording is as follows:

> I would like to get your feelings toward some of our political leaders and other people who are in the news these days. I will read the name of a person, and I would like you to rate that person using the feeling thermometer. Ratings between 50 and 100 mean that you feel favorably and warm toward the person; ratings between 0 and 50 degrees mean that you do not feel favorably toward the person and that you do not care too much for that person. You would rate the person at the 50-degree mark if you do not feel particularly warm or cold toward the person. If we come to a person whose name you do not recognize, you do not need to rate that person. Just tell me, and we will move on to the next one.

The thermometers thus can be treated as a continuous variable with 0 standing for *very much dislike* and 100 for *like very much*. Notice that the question wording invites respondents to consider qualitative distinctions within the thermometer spectrum: 0–49 scores indicate cold

feelings, 50 indicates a lukewarm affection towards the object, and 51–100 indicate warm feelings (Anderson & Granberg, 1991). Consequently, we use two versions of leader thermometers. One is a continuous variable (divided by 10) which ranges from 0 (*dislike*) to 10 (*like*), and the second is a categorical variable with three categories for each affection: *dislike* (0–49), *lukewarm* (50), and *like* (51–100).

The second independent variable is partisanship. This variable was originally measured on a seven-point scale, in which respondents were asked about their partisanship identification. The responses on this scale included: strong Democrat, weak Democrat, independent–Democrat, independent, independent–Republican, weak Republican, and strong Republican. In most analyses, we distinguish between partisans, either strong, weak, or leaner, and independents. Some of the analyses focus on in-party and out-party leader evaluations, excluding independents from the analysis. A combination of partisanship and the thermometers includes an in-party leader evaluation, i.e., the thermometer score of the party a respondent identifies with, and an out-party leader evaluation, i.e., the thermometer score of the party with which a respondent does not identify. Independents are excluded from these measures as their in-/out-party affections are indeterminate.

For reasons already outlined, the models are tested using a standard set of demographic controls, namely, age, gender, race, college degree, and income. The ANES data do not include information about past voting in some of the years, so we do not include them in the analysis here. As it happens, the inclusion or exclusion of this variable has no discernible impact on the main findings (see Tables A6 to A9 in the Supplementary File). An equalizing weight is added so that each survey counts equally in the pooled estimation. The dependent variable is Voter Turnout which is coded as 1 if the respondent voted in the last elections and 0 if she did not vote. As with other survey data, voter turnout is nearly always over-reported (Cassel, 2003); therefore, turnout rates reported in the empirical analysis are mostly higher than official turnout data in the US.

3.2. A Matrix of Affections Towards the Two Candidates and Its Effect on Turnout

We begin with a descriptive analysis of candidate thermometers over time. Recall that the thermometer scales can be recoded into three categories: like, lukewarm, and dislike. When applied to the Democratic and Republican candidates, that coding strategy produces six combinations of the three types of affection towards each of the two candidates. In effect, each voter can be assigned to one of the six following options:

1. Lukewarm towards both candidates;
2. Dislike one candidate and be lukewarm towards the other;
3. Like one candidate and be lukewarm towards the other;
4. Like one candidate and dislike the other
5. Like both candidates;
6. Dislike both candidates.

The distributions of cases across those six categories for the period 1968–2020 are as follows: Almost 12% of the voters do not like any of the presidential candidates. They either dislike both candidates (Category 6, 4.5%), dislike one and are lukewarm towards the other (Category 2, 3.4%), or are lukewarm towards both candidates (Category 1, 3.8%). Thus, 88% of voters hold positive affection towards at least one candidate. The largest group is of voters that like one candidate and dislike the other (Category 4, 53%); 14.5% like one and are lukewarm towards the other (Category 3), and the last group of voters likes both candidates (Category 5, 21% of respondents).

Before delving into the effect of these categories on turnout per H1, the place to begin is with an overview of the distribution of these categories among the American electorate over time. This exploration speaks to the broader question of increasing polarization in US politics, not from a partisan perspective but rather as reflected in evaluations of presidential candidates. We also compare the distribution of candidate evaluations between partisans—Democratic and Republicans—and independents. Figure 1 illustrates the distribution of the six categories in each presidential election year from 1968 to 2020 (Figure 1, Panel A) and the pooled data for Democrats, Republicans, and independents (Figure 1, Panel B). The cross-time comparisons reveal some variance in the distributions. There is a clear trend in the data: The proportion of voters who like both presidential candidates has declined. At the same time, the group of voters who like one candidate and dislikes the other—those with a clear-cut choice—has grown. This trend seems to signify the growing affective polarization among the American electorate (Iyengar et al., 2019), but it may also reflect the polarizing effect of Trump's candidacies.

When examining the groups by partisanship (Figure 1, Panel B), the graph indicates that the differences between Democratic and Republican partisans are modest. Those who identify as Republicans are less likely to report being lukewarm towards either candidate compared to Democrats, and they are more likely than Democrats to like one candidate and dislike the other. Independents (the middle bar) are much more likely to express lukewarm feelings towards either candidate (33% report feeling lukewarm towards at least one candidate). They are also more inclined than partisans to either like or dislike both candidates. Another difference between independents and partisans is the size of the group of respondents who sharply differentiate between candidates by liking one candidate and disliking the other candidate. This group constitutes more than

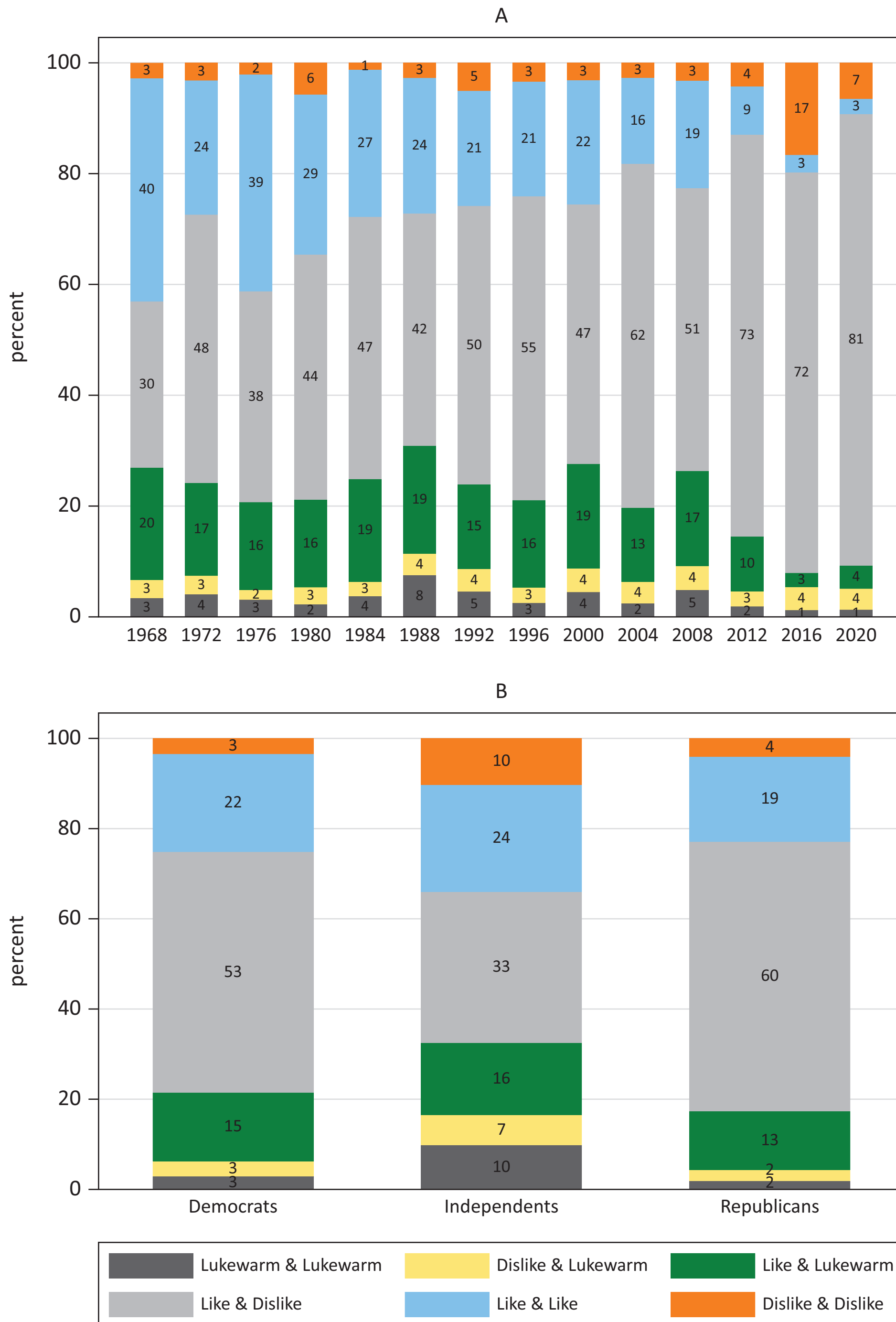

Figure 1. Distribution of coded candidate thermometers across election years (Panel A) and by partisanship (Panel B). Source: Authors' work based on ANES 1968–2020.

half of the respondents among partisans, while among independents, it is much smaller and constitutes 33% of respondents.

The central empirical question to consider is: What effects do these candidate evaluations have on levels of voter turnout? These effects are estimated using a logistic regression model with the turnout as a dependent variable and the six groups of candidate affections described above as independent dummy variables (with the fourth group, Like & Dislike, as the reference category). Controls for age, gender, race, education, and income, as well as election fixed effects, are included. The results are reported in Table A1 in the Supplementary File. The coefficients for each group are presented in Figure 2. In accordance with H1, the group of respondents who most differentiate between candidates, i.e., those who like one candidate and dislike the other, reports the highest propensity to vote. Their calculated probability of voting is 0.84. The two other large groups of voters, those that either like the two

candidates or like one and are lukewarm towards the other, have somewhat lower vote probabilities: 0.79 and 0.76, respectively. Note that, when examining the effect among the other groups, who comprise alienated and/or indifferent voters, those who dislike both candidates or dislike one candidate and are lukewarm towards the other are more likely to vote (0.75 and 0.72, respectively) than those who feel lukewarm towards both candidates and qualify as a completely indifferent group of voters (0.60). In effect, even citizens expressing complete negativity towards the candidates are more likely to vote than the ones who report no feelings whatsoever. To be sure, these groups comprise a very small portion of the American electorate.

While our analyses do not aim to explain overall turnout levels (but rather the impact of leader evaluations on individual citizens), there is some evidence of an association between the distribution of leader evaluations among ANES respondents in each year and the reported turnout levels. The strongest correlation is between the share of respondents with distinctive evaluations, those who like one candidate and dislike the other, and overall turnout levels. The higher the number of respondents with this structure of evaluations, the higher the overall reported turnout in that election. Table A5 in the Supplementary File presents the full results.

3.3. Liking "Your" Candidate or Disliking the "Opponent": What Mobilizes Voters?

So far, the analysis shows that citizens who differentiate between the presidential candidates are the most likely to vote, and citizens with lukewarm feelings towards both candidates are the least likely to vote. But what the preceding analysis has not considered is the possibility that some voters might hold party affiliations that are at odds with their candidate evaluations. Among the respondents who like one candidate and dislike the other, 84% like their in-party candidate and dislike the out-party candidate, 8% dislike the in-party candidate and like the other candidate, and 8% qualify as pure independents. When combined with turnout, respondents whose affections align with party identification are more likely to vote compared to the ones with candidate evaluations that are at odds with their party identification. The implication clearly is that party identification matters for leader evaluations and that considering evaluations without taking into account partisanship ignores the possible impact on turnout of dissonance between party affinity and evaluation of the party's current leader. Furthermore, the effect of candidate evaluations may be contingent not only on in-party evaluations but also on the evaluations of the out-party candidate. Citizens who dislike their own party's candidate are more likely to vote if they happen to like the out-party candidate, compared to voters who dislike her.

What needs to be explored, then, is the effect of the voter's feelings towards her in-party candidate on turnout while controlling for her feelings towards the out-party candidate. The place to begin is by considering the cross-time effect of candidate thermometers on turnout by partisanship. In this case, the analysis is based on a regression model of the pooled ANES data for the years 1968–2020. It estimates the interaction between

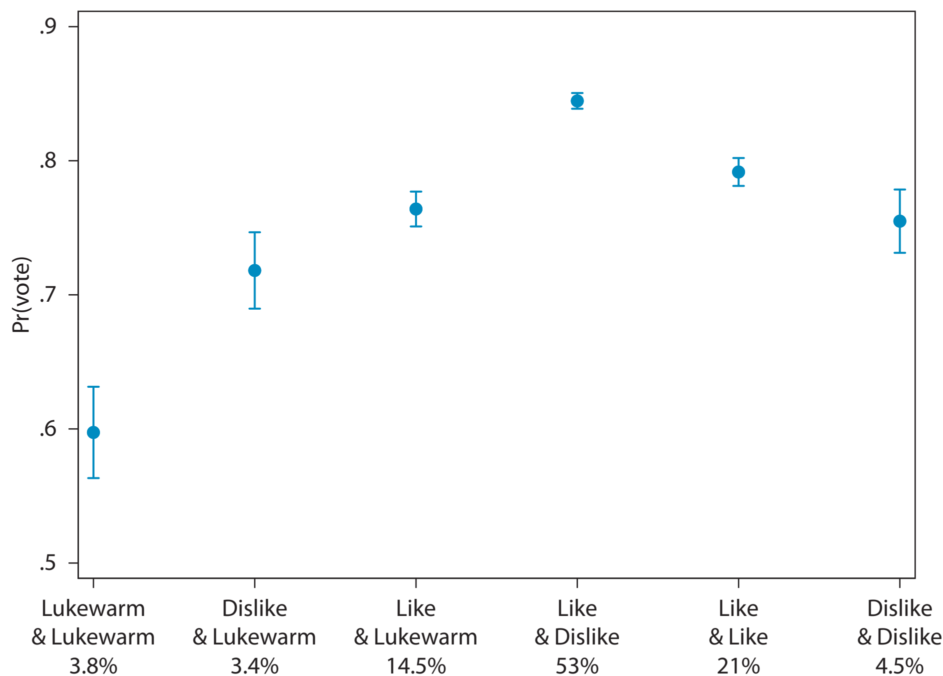

Figure 2. The effect of candidate evaluations on turnout. Notes: Predicted probabilities of voting for six different combinations of candidate evaluations; results are based on Table A1 of the Supplementary File; the labels on the horizontal axis indicate the share of each of the six groups in the data. Source: Authors' work based on ANES 1968–2020.

candidate thermometer and election years. The analysis is conducted separately for Democrats and Republicans. The models include controls for the out-party candidate thermometers as well as for age, gender, race, education, and income, as well as election fixed effects. Full results are reported in Table A2 in the Supplementary File. Figure 3 presents the effects of candidate thermometers on their in-party voters (Panels A1 and A2 for Democrats and Panels B1 and B2 for Republicans).

Panels A1 and B1 show that affection towards the in-party candidate has an impact on Democratic and Republican partisans in the predicted direction: The more they like the in-party candidate, the more likely they are to vote, holding the evaluation towards the out-party candidate constant. However, this effect is not significant in all election years. For example, notice that positive feelings towards Obama in 2008 and 2012 significantly mobilized Democrat voters. For Republicans, positive affection towards George H. W. Bush in 1988 had a significant effect on voting. Of course, the obverse also holds: The more voters dislike the candidate of their party, the less likely they are to vote. The results support H2a, arguing that positive feelings mobilize voters while negative feelings demobilize voters, although the causal direction is not settled.

What about affection towards the out-party candidate? Panels A2 and B2 report the effect of thermometer scores for out-party candidates holding the evaluation towards the in-party candidate constant. The expectation per H2b is that liking the opposing candidate introduces dissonance with a corresponding demobilizing effect. Citizens might prefer to stay at home rather than support the candidate of the other party. Indeed, the picture that emerges is of mostly negative effects. But, not surprisingly, there is some variation between election years. For example, the more Democrats liked George W. Bush in 2004 or Trump in 2016 and 2020, the less likely they were to vote. For Republicans, liking Bill Clinton in 1992 had a negative effect on voting. The "personalization of politics" thesis suggests that candidate thermometers should matter more now than in the past. Our data, however, do not endorse that straightforward expectation. The impact of leader evaluations on turnout, if anything, appears to be relatively stable. To be sure, there are variations, but those variations might be better described as election specific.

While the effect of candidate thermometers on turnout among partisans is rather clear, for those who identify as independents, Figure 4 shows the pattern is much less consistent. In some elections, the

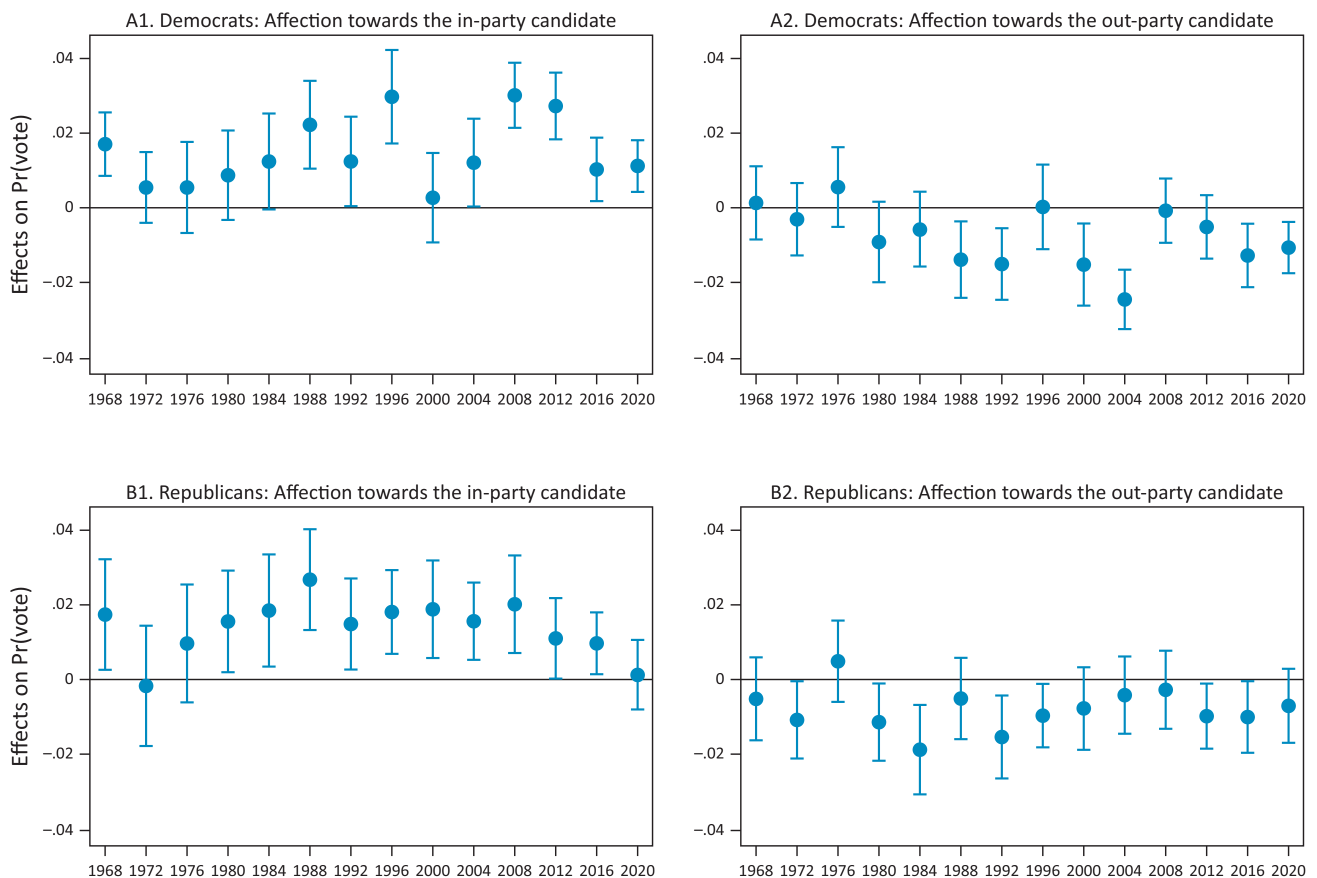

Figure 3. The effect of affection towards presidential candidates on turnout, 1968–2020. Notes: Panels A1 and A2 present the results for Democratic voters; Panels B1 and B2 present the results for Republican voters; results are based on Models 1 and 2 in Table A2 of the Supplementary File. Source: Authors' work based on ANES 1968–2020.

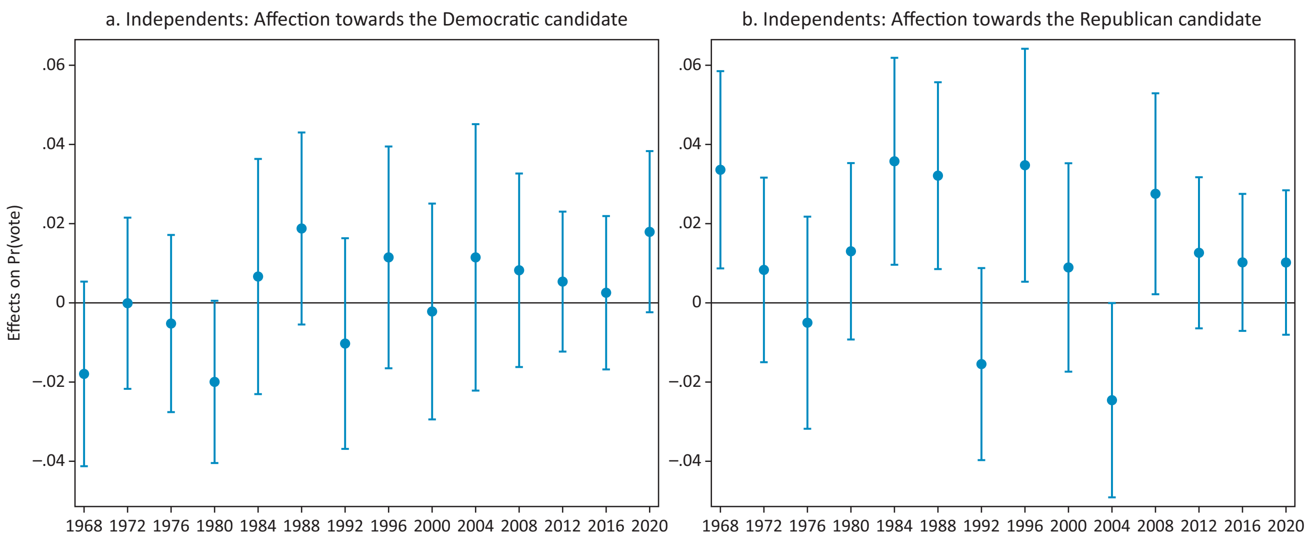

Figure 4. The effect of affection towards presidential candidates on turnout among independents: (a) Democratic candidates and (b) Republican candidates. Note: Results are based on Model 3 in Table A2 of the Supplementary File. Source: Authors' work based on ANES 1968–2020.

thermometer scales have a positive effect on turnout, but in others there is a negative effect. It is noteworthy that these effects tend to be significant for the Republican thermometers but not for the Democratic thermometers. For example, higher scores of the Republican thermometers increased turnout in 1984 and 1988, but in 2004 they decreased the turnout level among independents. This implies that non-partisan citizens are mobilized or demobilized by leader evaluations in a way that varies across elections.

3.4. Do Negative Feelings for Out-Party Candidates Mobilize Voters?

The concluding section of the analysis turns to the question: Do negative feelings towards the out-party candidate mobilize voters who like their in-party candidate and voters who dislike their in-party candidate to the same degree? Indeed, the motivation to prevent a disliked candidate from winning might be the same for all. Nonetheless, it could be that this motivation applies only to voters who also have a candidate for whom they would like to vote. The impact of the interacted effect of in-party candidate and out-party candidate evaluations is evaluated with a regression model that includes the interaction between these two variables as well as demographic control variables and election fixed effects. For presentation purposes, the out-party thermometer is reversed so that higher values signify disliking the out-party candidate. Results can be found in Table A3 in the Supplementary File. The following analyses (Figures 5 and 6) are performed for Democrats and Republicans together. Separate models yield overall similar results for each group. Figure 5 presents the effect of the out-party thermometer on turnout as a function of the affection towards the in-party candidate. Here, it is clear that disliking the out-party candidate has a positive effect on turnout, but only when respondents have positive feelings towards their in-party candidate. For those who dislike the in-party candidate, the out-party candidate has no impact on turnout. Moreover, among those who completely dislike "their" candidate, negative feelings towards the other candidate can even have a demobilizing effect, probably indicating their general dissatisfaction with both candidates. These results support H3.

One way to evaluate these findings in an even more nuanced way is to re-run a similar model but use the ordinal measures of the three-categories measure of the thermometer scales (like, lukewarm, and dislike). The results of that approach are reported in Table A4 in the Supplementary File. Panel A in Figure 6 presents the probability of voting as a function of the respondent's affection (dislike, like, or lukewarm) towards the in-party leader (the x-axis). For each in-party effect, the graph presents the probability of voting in relation to the respondent's affection towards the out-party leader. The circles represent respondents who dislike the out-party leader. The squares stand for respondents who are indifferent towards the out-party leaders, while the diamonds represent respondents who like the out-party leader.

The graph shows that among respondents who like the in-party candidate (the right category on the x-axis), those who dislike the out-party candidate exhibit the highest predicted probability to vote: 0.86. Thus, disliking the out-party candidate mobilizes voters compared to those who either like or are lukewarm towards the other candidate. The probability of those groups voting is 0.76 and 0.74, respectively. In effect, for respondents who like their candidate, negative feelings towards the opponent can increase their likelihood of voting by about 10 percentage points compared to others with lukewarm or positive feelings towards the out-party candidate.

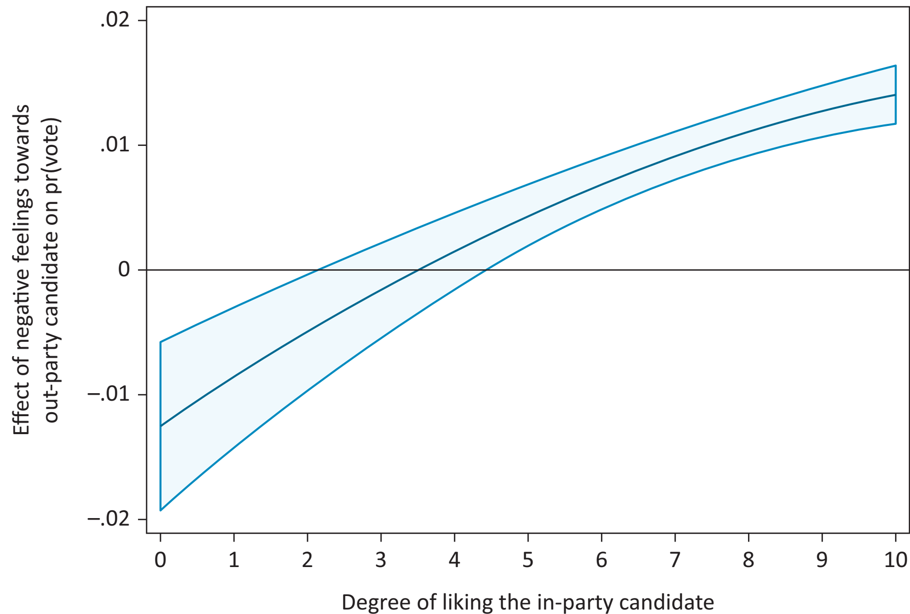

Figure 5. The interacted effect of in-/out-party candidate affection on turnout. Notes: The graph presents the marginal effect of negative feelings towards the out-party candidate on turnout as a function of the in-party candidates' thermometer; results are based on Table A3 in the Supplementary File. Source: Authors' work based on ANES 1968–2020.

What about voters who either dislike or are lukewarm towards the in-party candidate? In these cases, the results show that negative feelings do not have the same mobilizing effect. For both groups, it does not matter whether the voter likes or dislikes the other candidate. Liking or disliking under these conditions yields similar predicted vote probabilities. Thus, voters who do not like their in-party candidate will be more likely to vote if they like or dislike the other candidate. The significant finding here concerns the importance of having some affection, either positive or negative, for mobilizing the vote. When partisans do not like the current leader of their party, they nonetheless will be motivated to vote as long as they have a certain affection towards the candidate of the other party. Having lukewarm feelings is related to lower levels of turnout. But both positive and negative affects towards leaders increase the probability of voting, even if they are directed towards the candidate of the party with which the voter does not identify.

Panel B in Figure 6 presents the other side of the interaction, namely, the degree to which in-party candidate evaluation depends on the out-party candidate evaluation. Here the data show that liking the in-party candidate is associated with a higher probability of voting. Yet, among respondents who like the out-party candidate, liking or disliking the in-party candidate yields

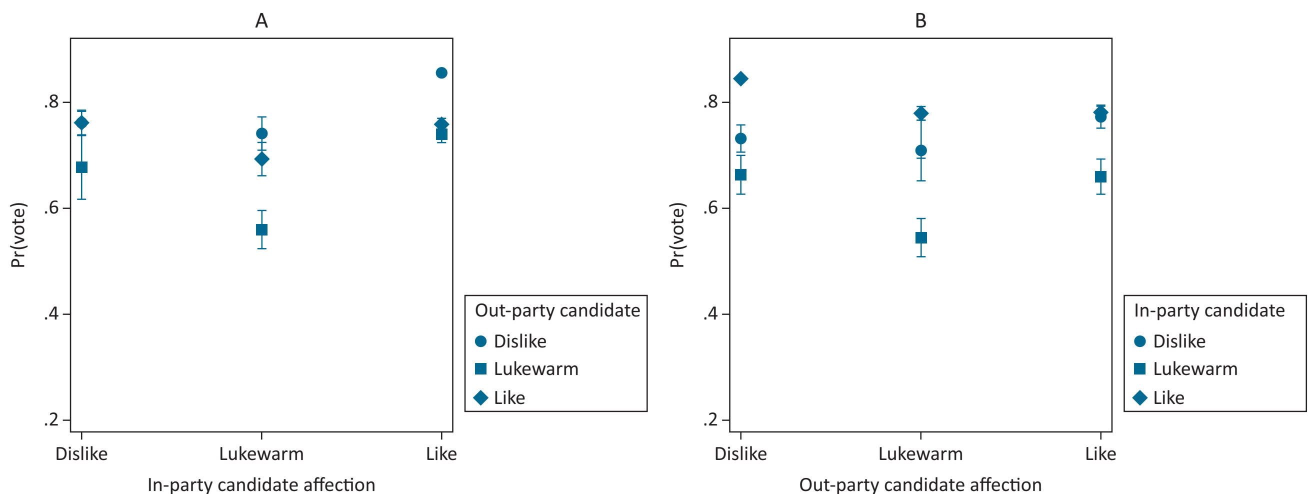

Figure 6. The interacted effect of in-/out-party candidate affection on turnout by categories. Note: The graph presents predicted probabilities to vote by leader evaluations based on the interacted effects specified in Table A4 in the Supplementary File. Source: Authors' work based on ANES 1968–2020.

similar levels of turnout. In effect, these results show that having a clear-cut choice between candidates who run counter to your party affiliation leads to similar levels of turnout as voters who hold positive affections towards both candidates.

4. Concluding Discussion

With the increasing focus on politicians and party leaders in contemporary politics, scholars of voting behavior have been trying to identify the various impacts that these political actors might have on citizens' political behavior. There is ample evidence that party leaders affect voters' decisions in terms of vote choice. The degree to which turnout can be affected by voters' feelings towards party leaders is underexplored, particularly so given speculations that leaders matter more now than before to electoral outcomes. Earlier investigations yielded some important insights into how leader-voter turnout dynamics might work (Brody & Grofman, 1982; Weisberg & Grofman, 1981), but those investigations were "early" in the sense that they predated the growing embrace of the presidentialization hypothesis. Moreover, their empirical findings were somewhat inconclusive. This investigation has revisited the possible connections between leader evaluations and voter turnout, and it has done so across a much greater time span using more than 50 years' worth of ANES data.

The findings reported here show that leader evaluations unequivocally do matter to levels of voter turnout in recent presidential elections in the US. But they do so in somewhat nuanced ways. We demonstrate that leader evaluations and citizens' turnout is mediated by party identification. More particularly, the evidence is that the likelihood of voting is affected by (a) the degree to which voters' affections towards the candidates differ from one another and (b) the extent of congruence between party affiliation and the voter's affections towards the presidential candidates of both parties. First, the data show that respondents who express clear preferences, that is, those who hold positive feelings towards one candidate and negative feelings towards the other, have a higher probability of voting than other voters. That is in stark contrast to voters who express no definite feelings (positive or negative) towards both candidates. Not surprisingly, perhaps, Republicans and Democrats are more likely to vote when they like their in-party candidate for the presidency. Conversely, voters of both groups are *less* likely to vote when they like the out-party candidate. That dissonance depresses voter turnout. These effects, however, are not entirely uniform; they vary in their impact across elections. And it is noteworthy that the effects are somewhat asymmetrical. Independents are more affected by their affection towards the Republican candidate, while thermometers for the Democratic candidate tend not to affect independents' likelihood of voting. That finding warrants deeper investigation. This is so not least of all because more Americans claim to be independent. Lastly, our data reveal another noteworthy asymmetry, namely, the assumed mobilization effect among those who hold negative feelings towards the out-party candidate operates only one under one condition, namely when one likes the in-party candidate. This finding speaks to the debate on negativity and voting (Martin, 2004; Nai, 2013). Negative feelings can mobilize voters, but campaigners should be careful not to completely rely on negativity towards the other side. That calculus ignores the critical role played by voters' evaluations of the in-party candidate.

The case of the US brings with it a number of analytical advantages for investigating links between voter turnout and leader evaluations. First, the presence of a two-party system presents voters with a relatively straightforward choice set. Second, the exceptional durability of that two-party system both underpins a correspondingly consistent foundation for patterns of party identification and diminishes the likelihood that cross-time variations in voter turnout could be reasonably attributed to changes in electoral arrangements or the party system. A third advantage flows from the character of the long-running ANES itself. Patterns of stability and change are more reliably discerned with data collected over a longer time span. Equally significant, the ANES has been strikingly consistent in using the very same measures of such variables as party identification, voter turnout, and leader evaluations, which have been central to the preceding analysis. Consequently, it is unlikely that the observed variations could simply be dismissed as functions of instrumentation effects. Together, these attributes increase confidence that the data findings are robust. If the US qualifies as exceptional in these respects, then the reported findings raise other research questions: Do the same leader-voter turnout dynamics apply in other settings? Do variations in those dynamics correspond to different regime styles, electoral arrangements, or party systems? Analysts have made important recent contributions in applying a similar line of analysis to the European setting (da Silva, 2018; da Silva & Costa, 2019; da Silva et al., 2021). What is called for now is a more expansive research effort to determine what are the key system-level characteristics that gear those relationships.

Acknowledgments

We thank Chitralekha Basu, the editors, and the reviewers for their very helpful comments and suggestions. We are also grateful to Yaniv Shapira for the excellent research assistance. An earlier version of this article was presented at the annual meeting of the Midwest Political Science Association (2018).

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Aarts, K., Blais, A., & Schmitt, H. (Eds.). (2011). *Political leaders and democratic elections*. Oxford University Press.

Adams, J., Dow, J., & Merrill, S. (2006). The political consequences of alienation-based and indifference based voter abstention: Applications to presidential elections. *Political Behavior*, *28*(1), 65–86.

Anderson, E., & Granberg, D. (1991). Types of affective evaluators in recent US presidential elections. *Polity*, *24*(1), 147–155.

Balmas, M., Rahat, G., Sheafer, T., & Shenhav, S. R. (2014). Two routes to personalized politics: Centralized and decentralized personalization. *Party Politics*, *20*(1), 37–51.

Bankert, A. (2021). Negative and positive partisanship in the 2016 U.S. presidential elections. *Political Behavior*, *43*(4), 1467–1485.

Bankert, A. (2022). Negative partisanship among Independents in the 2020 US presidential elections. *Electoral Studies*, *78*, Article 102490.

Barisione, M. (2009). So, what difference do leaders make? Candidates' images and the "conditionality" of leader effects on voting. *Journal of Elections, Public Opinion and Parties*, *19*(4), 473–500.

Bartels, L. M. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science*, *44*(1), 35–50.

Blais, A. (2006). What affects voter turnout? *Annual Review of Political Science*, *9*(1), 111–125.

Blais, A., & Carty, R. K. (1990). Does proportional representation foster voter turnout? *European Journal of Political Research*, *18*(2), 167–181.

Brady, H. E., Verba, S., & Scholzman, K. (1995). Beyond SES: A resource model of political participation. *American Political Science Review*, *89*(2), 271–295.

Brody, R. A., & Grofman, B. (1982). Stimulus differentiation versus stimulus complexity as factors affecting turnout in two-candidate and multicandidate races. *Political Behavior*, *4*(1), 83–92.

Brody, R. A., & Page, B. I. (1973). Indifference, alientation and rational decisions. *Public Choice*, *15*(1), 1–17.

Cassel, C. A. (2003). Overreporting and electoral participation research. *American Politics Research*, *31*(1), 81–92.

Converse, P. E. (1969). Of time and partisan stability. *Comparative Political Studies*, *2*(2), 139–171.

da Silva, F. F. (2018). Fostering turnout? Assessing party leaders' capacity to mobilize voters. *Electoral Studies*, *56*, 61–79.

da Silva, F. F., & Costa, P. (2019). Do we need warm leaders? Exploratory study of the role of voter evaluations of leaders' traits on turnout in seven European countries. *European Journal of Political Research*, *58*(1), 117–140.

da Silva, F. F., Garzia, D., & De Angelis, A. (2021). From party to leader mobilization? The personalization of voter turnout. *Party Politics*, *27*(2), 220–233.

Downs, A. (1957). *An economic theory of democracy*. Harper & Row.

Elgie, R., & Passarelli, G. (2018). Presidentialisation: One term, two uses between deductive exercise and grand historical narrative. *Political Studies Review*, *17*(2), 115–123.

Fiorina, M. P. (1981). *Retrospective voting in American national elections*. Yale University Press.

Fowler, J. H. (2006). Habitual voting and behavioral turnout. *The Journal of Politics*, *68*(2), 335–344.

Franklin, M. N. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945*. Cambridge University Press.

Garzia, D. (2012). Party and leader effects in parliamentary elections: Towards a reassessment. *Politics*, *32*(3), 175–185.

Garzia, D., & da Silva, F. F. (2021). Negative personalization and voting behavior in 14 parliamentary democracies, 1961–2018. *Electoral Studies*, *71*, Article 102300.

Gattermann, K., & de Vreese, C. H. (2022). Understanding leader evaluations in European Parliament elections. *European Union Politics*, *23*(1), 141–160.

Geys, B. (2006). Explaining voter turnout: A review of aggregate-level research. *Electoral Studies*, *25*(4), 637–663.

Gidengil, E. (2011). Voter characteristics and leader effects. In K. Aarts, A. Blais, & H. Schmitt (Eds.), *Political leaders and democratic elections* (pp. 147–164). Oxford University Press.

Glenn, N., & Grimes, M. (1968). Aging, voting and political interest. *American Sociological Review*, *33*(4), 563–575.

Huddy, L., & Bankert, A. (2017). Political partisanship as a social identity. In W. R. Thompson (Ed.), *Oxford research encyclopedia of politics*. https://doi.org/10.1093/acrefore/9780190228637.013.250

Inglehart, R., Norris, P., & Ronald, I. (2003). *Rising tide: Gender equality and cultural change around the world*. Cambridge University Press.

Iyengar, S., & Krupenkin, M. (2018). The strengthening of partisan affect. *Political Psychology*, *39*(S1), 201–218.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*, 129–146.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405–431.

Jackman, R. W. (1987). Political institutions and voter turnout in the industrial democracies. *The American Political Science Review*, *81*(2), 405–424.

Jackman, R. W., & Miller, R. A. (1995). Voter turnout in the industrial democracies during the 1980s. *Comparative Political Studies*, *27*(4), 467–492.

Lavine, H. G., Johnston, C. D., & Steenbergen, M. R. (2012). *The ambivalent partisan: How critical loyalty promotes democracy*. Oxford University Press.

Lefkofridi, Z., Giger, N., & Gallego, A. (2014). Electoral participation in pursuit of policy representation: Ideological congruence and voter turnout. *Journal of Elections, Public Opinion and Parties*, *24*(3), 291–311.

Lobo, M. C. (2014). Party dealignment and leader effects. In M. C. Lobo & J. Curtice (Eds.), *Personality politics? The role of leader evaluations in democratic elections* (Vol. 7, pp. 148–166). Oxford University Press.

Martin, P. S. (2004). Inside the black box of negative campaign effects: Three reasons why negative campaigns mobilize. *Political Psychology*, *25*(4), 545–562.

Mughan, A. (2000). *Media and the presidentialization of parliamentary elections*. Palgrave.

Mughan, A. (2009). Partisan dealignment, party attachments and leader effects. *Journal of Elections, Public Opinion and Parties*, *19*(4), 413–431.

Nai, A. (2013). What really matters is which camp goes dirty: Differential effects of negative campaigning on turnout during Swiss federal ballots. *European Journal of Political Research*, *52*(1), 44–70.

Norris, P. (1999, October 28–29). *Blaming the messenger? Political communications and turnout in EU elections* [Paper presentation]. Citizen Participation in European Politics, Stockholm, Sweden.

Norton, P. (2003). The presidentialization of British politics. *Government and Opposition*, 38, 274–278.

Poguntke, T., & Webb, P. (Eds.). (2005). *The presidentialization of politics: A comparative study of modern democracies*. Oxford University Press.

Powell, G. B. (1986). Voter turnout in comparative perspective. *American Political Science Review*, *80*(1), 17–43.

Radcliff, B., & Davis, P. (2000). Labor organization and electoral participation in industrial democracies. *American Journal of Political Science*, *44*(1), 132–141.

Rahat, G., & Kenig, O. (2018). *From party politics to personalized politics? Party change and political personalization in democracies*. Oxford University Press.

Samuels, D. J., & Shugart, M. S. (2010). *Presidents, parties, and prime ministers: How the separation of powers affects party organization and behavior*. Cambridge University Press.

Simas, E. N., & Ozer, A. L. (2021). Polarization, candidate positioning, and political participation in the US. *Electoral Studies*, *73*, Article 102370.

Smets, K., & van Ham, C. (2013). The embarrassment of riches? A meta-analysis of individual-level research on voter turnout. *Electoral Studies*, *32*(2), 344–359.

Stockemer, D. (2016). Is the turnout function in democracies and nondemocracies alike or different? *Politics & Policy*, *44*(5), 889–915.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178.

Verba, S., Nie, N. H., & Kim, J. O. (1978). *Participation and political equality: A seven-nation comparison*. Cambridge University Press.

Weisberg, H. F., & Grofman, B. (1981). Candidate evaluations and turnout. *American Politics Quarterly*, *9*(2), 197–219.

About the Authors

Liran Harsgor is a lecturer in the Division of Government and Political Theory, School of Political Science, University of Haifa. Her research focuses on cross-national and comparative analysis of political attitudes, policy preferences, and voting behavior. In the past, she has been a Halbert post-doctoral fellow at the Munk School of Global Affairs and Public Policy, University of Toronto (2017–2018).

Neil Nevitte is professor emeritus of the Department of Political Science and Munk School of Global Affairs and Public Policy, University of Toronto. As co-investigator of four federal election studies and principal investigator of the World Values Surveys (Canada), his research focuses mostly on elections and public opinion. His publications have appeared in such journals as *Comparative Political Studies*, *British Journal of Political Science*, and the *Journal of Democracy*. He has also authored or co-authored 10 book-length projects concerned with elections or value change.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 374–383
https://doi.org/10.17645/pag.v10i4.5752

Article

Loud and Negative: Exploring Negativity in Voter Thoughts About Women and Men Politicians

Tobias Rohrbach [1,2]

[1] Department of Media and Communication Research, University of Fribourg, Switzerland; tobias.rohrbach@unifr.ch
[2] Department of Political Science, University of Amsterdam, The Netherlands

Submitted: 2 May 2022 | Accepted: 4 July 2022 | Published: 30 December 2022

Abstract
Negative information about political candidates is readily available in contemporary political communication. Moreover, negativity is tightly connected to gendered expectations about what constitutes appropriate behavior for politicians. Yet, existing theoretical models of negativity and candidate evaluation typically do not address the role of gender and the available empirical evidence remains inconclusive regarding the electoral consequences of the interaction of negativity and gender. This article tackles these gaps in two studies to investigate how negativity manifests in voters' thoughts about women and men politicians in response to negative media cues and how these thoughts affect vote preference. Study 1 uses a mixed methods think-aloud approach to trace the first impression formation and subsequent decision-making process (N = 78). Study 2 replicates the design as an online thought listing survey experiment (N = 142). A similar quantitative pattern emerges across both studies: (a) Negative cues elicit similar amounts of negativity in voters' thoughts for women and men politicians, (b) these negative thoughts strongly lower candidates' electoral chances, (c) but less so for women candidates. The qualitative analysis suggests that negative cues heuristically affect earlier stages of impression formation while voters are likely to rely on gender cues when they rationalize their vote decision.

Keywords
candidate evaluation; gender; negativity; think aloud; thought listing

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

More women are entering politics than ever before (Hughes & Paxton, 2019) and there are indeed signs that the political tide is turning in women's favor. Recent research suggests that voters display little bias against women candidates at the ballot box (e.g., Bridgewater & Nagel, 2020; Dolan, 2014) or in experimental settings (Schwarz & Coppock, 2022). At the same time, negative political campaigns are commonplace (Nai, 2020) and may be becoming more frequent (Geer, 2012; for a meta-analysis see Lau et al., 2007). An interplay of structural, contextual, and personal factors is driving political candidates to incorporate elements of negativity in their campaign strategies (e.g., Valli & Nai, 2020), which drive voters' attention directly (Fridkin & Kenney, 2012) or indirectly by generating more (negative) media coverage (Maier & Nai, 2020; Meffert et al., 2006; Soroka et al., 2019).

These larger phenomena in political communication do not happen in isolation but affect the ways that voters evaluate political candidates in concert. Although existing models of negativity and candidate evaluation do not address the role of candidate gender (Klein & Ahluwalia, 2005; Lodge & Taber, 2013), negativity is tightly connected to gendered expectations about what constitutes appropriate behavior for politicians (Krupnikov & Bauer, 2014). The literature on gender stereotyping and candidate evaluation offers three possible theoretical

explanations for the interaction of negativity and candidate gender on voter evaluations.

First and most conventionally, a *reinforcing effect* of campaign negativity would be predicted by the research on role congruity (Eagly & Karau, 2002) and stereotype activation (e.g., Bauer, 2015). Many forms of negative politics—such as candidate attacks, mudslinging, or scandalization (Craig & Rippere, 2016; Fridkin & Kenney, 2012)—run counter to stereotypical expectations that prescribe women (but not men) to be warm, communal, and nurturing while proscribing any forms of aggressiveness, immorality, or stubbornness (Prentice & Carranza, 2002). Though voters may not *directly* punish women candidates on the basis of gender in neutral conditions (see Schwarz & Coppock, 2022), "campaign communication activates stereotypes when they otherwise might not be activated, thereby diminishing support for female candidates" (Bauer, 2015, p. 691). By reinforcing the perceived disconnect between leadership and gender stereotypes (Schneider & Bos, 2014), negativity in candidate messages—either in their own communication (see, e.g., Valli & Nai, 2020) or in media coverage (Van Der Pas & Aaldering, 2020)—can thus *indirectly* affect voter evaluations.

Second, however, an *equalizing effect* of campaign negativity may arise from voters' dislike of negative campaigning irrespective of the gender of the involved candidate (Fridkin & Kenney, 2012). In this logic, the attentional pull of negative cues in a candidate's message outweighs gender cues (Meffert et al., 2006; Soroka et al., 2019) and "neutralize[s] the disadvantages caused by gender stereotypes" (Gordon et al., 2003, p. 35). Similarly, research has shown that women who focus on masculine traits and issues in their communication can counteract gender stereotypes (Bauer, 2017). The underlying idea is that voters' decision to categorize a female candidate as either a political leader or a woman is not clear-cut but malleable by strategic communication. Messages containing stereotypically masculine forms of negativity (e.g., attacks, corruption, scandals, etc.) may thus shift the most salient category during the evaluation from "woman" to "leader" (Bauer, 2017) and provide voters with a way to ignore or reconcile incongruent role expectations by creating a new subtype (e.g., "female leader"; Schneider & Bos, 2014).

Third, the notion of benevolent sexism (Glick & Fiske, 1996) could also suggest a *protective effect* of campaign negativity for women. Negative campaign elements—especially those framing or portraying women as targets of attacks or scandals—violate the "norm of civility towards women" (Cassese & Holman, 2019, p. 57). In turn, this impression of exposed vulnerability for women candidates can compel voters to protect women by not only excusing or overcompensating for any potential transgression but also by punishing their (male) opponents (Barnes et al., 2020).

The mixed empirical evidence on the interaction of negative candidate messages and candidate gender does not clearly favor one theoretical argument over the other. In line with the reinforcement perspective, some studies indeed find that going negative on the campaign trail entails stronger backlash for women candidates than for men (Cassese & Holman, 2018; King & McConnell, 2003). Triangulating three different studies, Nai et al. (2021) have recently shown that voters consistently punish women—but not men—for using negative campaigning elements. In contrast, other research shows that negativity may act as an equalizing force resulting in few and inconsistent gender differences in candidate evaluations (Craig & Rippere, 2016; Krupnikov & Bauer, 2014). Finally, a few studies indicate in line with the protective perspective that the "presence of gender stereotypes appears to soften the blow of negative attacks" (Fridkin et al., 2009, p. 70; Gordon et al., 2003).

One explanation for these inconclusive findings might reside in the fact that reinforcing, equalizing, and protective effects are confounded by varying voter perceptions of negativity and that different forms (and definitions) of negativity may have different gendered consequences within and between voters (Sigelman & Kugler, 2003). Yet in-depth knowledge about the role of gender in voters' appraisal, processing, and application of negative information is still missing. I, therefore, propose to take a step back and approach the intersection of gender and negativity in an exploratory fashion. In two studies, I trace voters' thinking in response to negative and neutral candidate cues to assess differences in voters' thoughts about women and men candidates involved in negativity (RQ1) and to understand how voters' negative thoughts affect their vote decision (RQ2).

2. Study 1: Think Aloud Exploration

The think aloud (TA) paradigm conceptualizes the thinking process as a sequence of information chunks that enter participants' working memory for processing and verbalization while they perform a given task (Ericsson & Simon, 1993; Van Someren et al., 1994). While common in psychology and educational science (for a systematic review see Fox et al., 2011), the only use of concurrent verbalization techniques in the context of candidate evaluation—though not about gender or negativity—is a study by Lusk and Judd (1988), which traces voters' thoughts in response to candidate vignettes. The authors conclude that the strength of the TA method is to derive bottom-up perceptions of the investigated phenomenon, which addresses some of the definitory concerns of negativity (Sigelman & Kugler, 2003).

2.1. Participants

Seventy-two participants (51% women, M_{age} = 36.9, SD_{age} = 14.4) were recruited via snowball sampling by the author and nine research students of a master's seminar at a Swiss university. Participants had to be at least 18 years old and be fluent in German. The language

criterion served to ensure effortless verbalization as all materials were in German. In addition, special attention was paid to include participants with heterogeneous sociodemographic and professional backgrounds. The study protocol is part of a larger pre-registered project (available at https://osf.io/wgn9r) approved by the university's institutional review board.

2.2. Procedure and Materials

The TA paradigm follows a 2 (candidate gender: woman vs. man) × 3 (cue type: neutral vs. negative vs. unrelated) within-subjects quasi-experimental design consisting of a warm-up, two TA candidate evaluation tasks, and a brief post-test survey. First, participants were familiarized with the TA procedure in three rounds of warm-up tasks adapted from the TA literature (Ericsson & Simon, 1993; Van Someren et al., 1994).

The first evaluation task (T1) was designed to trace participants' initial responses to candidate cues. Each participant serially viewed mock newspaper title pages of fictional candidates, which are individually displayed for one minute each. Participants were instructed to spontaneously respond to the image and infer the candidates' political profile: "Please look at the image and try to guess what this person is like as a politician in real life."

For this task, a total of 14 candidate stimuli were grouped into seven different sets. Each set manipulated candidate gender (woman vs. man) and a specific framing by varying the title pages' candidate image and headline. A first set contained a neutral (pre-tested) portrait photo of a man or woman candidate with a headline simply identifying them as candidates for an election (see Panels A and B of Figure 1). Rather than focusing on one specific form of negativity, I manipulated three sets to explore different forms of negativity: (a) an image of a negative campaign ad denouncing a candidate as corrupt, (b) an image of a candidate displaying strong anger at a local debate, and (c) a paparazzi shot of a candidate at a strip club with a moralizing headline (see the Supplementary Materials). The remaining sets contained other framings and were used as filler materials for this study. Every participant viewed a total of six stimuli from three randomized sets (one neutral, one negative, and one filler). The design and content of all stimuli were adapted from real examples of media coverage and pilot tested.

The second task (T2) sought to capture participants' decision-making process. Participants were shown the same sets of title pages again but this time portraying the woman and man candidate simultaneously next to each other in the final stretch of a hypothetical race (see Panel C of Figure 1). Participants were instructed to make a choice: "Please look at the two candidates and think aloud about whom you would rather recommend to a friend." To mask the gender-specific intention of the study, participants viewed two neutral sets presenting same-gender races in addition to the sets from T1.

In the final part of the study, participants completed a short survey containing political, sociodemographic, and attitudinal measures.

2.3. Mixed Methods Analysis

2.3.1. Coding

The raw transcripts of the verbal report were first coded by means of qualitative content analysis. The annotated dataset was then used to extract measures for quantitative analysis (see Section 2.3.2). The individual candidate description (from T1) represented the unit of analysis. As each participant saw four candidate images (i.e., the neutral and one of three negative sets), this resulted in a total of 288 candidate descriptions. The verbal report for each candidate description was segmented into single thoughts as the unit of coding. Following recommended practice (Ericsson & Simon, 1993, pp. 172, 205–207, 266–270), a single thought was defined as a full sentence,

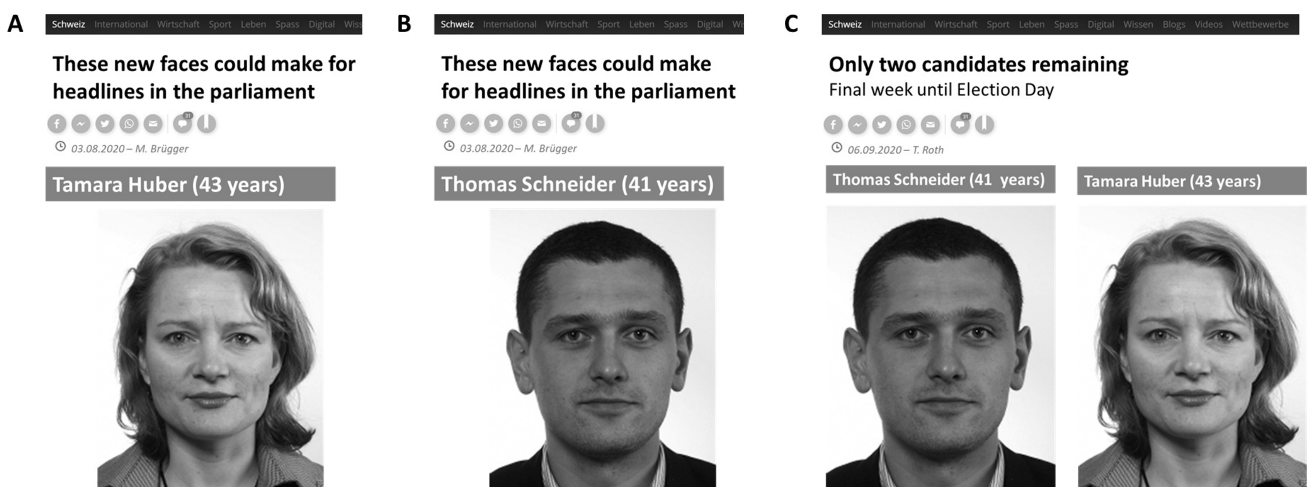

Figure 1. Neutral set of stimuli used in the first (Panel A and B) and second TA candidate evaluation task (Panel C). Note: Translated from German.

which represents the linguistic (and verbalizable) equivalent of a semantically closed unit of meaning.

In line with Lodge and Taber's (2013) dual process model of political evaluation, two dimensions of thinking were coded. First, *thought content* reflects the semantic core of the activated concept (i.e., what the thought is about). For this, a category grid was derived from the literature on candidate evaluation and inductively completed. The final category grid distinguished between six different thought contents (see Table 1 and the Supplementary Materials for the full category grid along with coding examples). Second, *thought affect*—that is, the general valence tendency accompanying the thought content—was categorically coded either as negative (−1), neutral or ambiguous (0), or positive (1).

All coding was conducted by the author and a student assistant after extensive training. In case of repeated inductively observed (sub-)categories or disagreements for coded thought contents, harmonizing decision rules were established and the material was revisited. Intercoder reliability for the more standardized thought affect was satisfactory (Krippendorff's α = 0.87).

2.3.2. Measures

The independent variables are *candidate gender* (0 = man, 1 = woman) and *cue type* (0 = neutral, 1 = negative) which are derived directly from the stimulus material.

As dependent variables, I measure *negativity* in voter responses as the sum of thoughts with negative affect per candidate description (M = 2.63, SD = 2.73) during T1. I capture participants' *vote choice* as a dummy variable to reflect whether they recommended the candidate (1) or not (0) during T2. Both measures are derived from the verbal report coding.

Finally, I include several individual characteristics as control variables. To account for differences in participants' verbalization speed, I measure their *total thoughts* as the sum of all verbalized thoughts per candidate coding in T1 (M = 7.13, SD = 2.65). From the short survey, I derive participants' own *gender* (0 = man, 1 = woman) and their *ideology* (two items on a scale from 1 = left/liberal to 10 = right/conservative; M = 3.06, SD = 1.14). Additionally, I assess *gender essentialism*, as gender essentialist beliefs can moderate the impact of gender stereotypes in candidate evaluations (Swigger & Meyer, 2019). Adapting their measure, I calculate an index of participants' average agreement to eight items (e.g., "Gender is a natural category") on a seven-point scale (M = 3.81, SD = 1.08).

2.3.3. Data Analysis

Data analysis simultaneously integrates quantitative and qualitative approaches where statistical analysis is used for identifying relationships and regularities and the qualitative in-depth analysis serves to explore and contrast underlying explanations (see Fearon & Laitin, 2013).

After a descriptive summary, I run Bayesian multilevel regression models to predict first the extent of negativity in participants' thoughts and then vote choice. I cluster the models around the stimulus set and the individual participant to accommodate the nested structure of the data. I rely on a Bayesian framework for its ability to draw conclusions based on probabilistic inferences about the presence—or absence—of an effect given the observed data (Gelman et al., 2013). Please refer to the Supplementary Materials for a technical discussion of model specification and evaluation.

I will report results as estimated posterior means along with 95% credible intervals (CrI). As a test of the evidence for or against the presence of an effect, I will calculate Bayes factors (BF). BF describe two models' predictive performance in relation to each other—that is, BF_{10} is calculated as the ratio of the likelihood of H_1 (evidence in favor of the presence of effect) over the likelihood of H_0 (evidence in favor the absence of effect)—given the data (Keysers et al., 2020; Wagenmakers et al., 2018). I follow the conventional classification for interpreting BF_{10} where a BF_{10} between 1 and 3 indicates anecdotal evidence, between 3 and 10 moderate evidence, between 10 and 30 strong evidence, between 30 and 100 very strong evidence, and a BF_{10} greater than 100 means extreme evidence in favor of the alternative hypothesis (see Hoijtink et al., 2019).

2.4. Results

Regarding the first research question, the pairwise comparisons show few systematic gender differences in negative thoughts about politicians (see Table 1). Across all thought content categories, participants have more negative thoughts about men (M = 3.30, SD = 3.17) than women candidates (M = 2.49, SD = 2.77). The BF_{10} for this comparison indicates that the presence of this small difference (d = 0.27) is 7.22 more likely than its absence. Note that this gender difference shrinks but persists in the multivariate analysis including control variables (BF_{10} = 3.96, see Model 1 in Table 2). No striking gendered patterns arise for single thought contents, except for candidates' personality, which is more frequently the object of negative thoughts for men than women politicians (BF_{10} = 5.83, d = 0.30).

Model 1 in Table 2 shows that negative candidate cues entail on average 1.31 (CrI = −1.32–3.85) more negative thoughts compared to neutral cues (BF_{10} = 6.69). The lack of evidence for an interaction effect indicates that negative cues increase negative thinking irrespective of candidate gender (BF_{10} = 2.51). Indeed, the qualitative analysis suggests that negative cues—regardless of the type of negativity—trigger negative thoughts across different thought content categories with no direct relation to the negative cue itself. For example, candidates involved in a scandal are not only criticized in terms of their integrity but also regarding their appearance, political experience, and competence. The following thought

Table 1. Summary of Bayesian t-tests comparing negativity in thoughts about women and men candidates across thought contents in Study 1.

	Man		Woman			
Thought content	*M*	*SD*	*M*	*SD*	*d*	BF_{10}
Layout and design	0.59	0.93	0.48	0.94	0.12	0.22
Candidate-related thoughts	2.26	2.41	1.75	2.24	0.22	2.41
Political profile	0.19	0.55	0.08	0.34	0.23	1.30
Personality Traits	0.89	1.29	0.56	0.95	0.30	5.83
Appearance	0.69	1.02	0.64	1.05	0.05	0.12
Personal life	0.25	0.75	0.16	0.52	0.14	0.28
Participant-related thoughts	0.45	0.98	0.25	0.64	0.24	1.45
All thoughts	3.30	3.17	2.49	2.77	0.27	7.22

passage of a male participant shows how an initial negative thought about the emotional display of the "angry candidate" can cascade into a stream of negativity of seemingly unrelated aspects:

> Oh wow, this guy looks pissed off, as if he wanted to bite off my head. He looks like the type of person who always shouts and never listens to distract from his incompetence. With this posture, he looks like a mulish bull. He has way too much gel in his hair and the way he holds up his chin makes me think of Mussolini. Very unlikeable. I now see that the image has almost no saturation, which makes it unpleasant to look at.

The second research question relates negativity in participant thoughts to their vote choice (see Model 2 in Table 2). The single most clear result is a negative effect of the number of negative thoughts on vote choice: With every additional negative thought, the chance of getting the participant's vote recommendation decreases by 37% on average (OR = 0.63, CrI = 0.49–0.77). The BF greater than 999 indicates extreme evidence.

Crucially, the effect of negativity on voter thoughts varies across candidate gender, with women being less strongly affected by participants' negative thoughts than men candidates (BF_{10} = 275.5). Panel A in Figure 2 illustrates this interaction and shows that men's chances of getting the vote drop dramatically when participants have only a few negative thoughts while the preference for women candidates diminishes much more gradually. The qualitative data point to a combination of equalizing and protective effects. For one, participants often struggle to form a decision after negative appraisals, calling their decision a "toss of a coin" (male, 31 years) or a "50–50 decision" (female, 54 years). In these cases, negativity appears to deflect from gender-related aspects and to pre-empt the potential of backlash against women candidates. Moreover, almost half the participants referred to the social context of structural bias against women when thinking about their vote

Table 2. Results of Bayesian multilevel regression models predicting negativity and vote choice in Study 1.

	Model 1: Negative Thoughts (T1)			Model 2: Vote Choice (T2)		
Predictors	Estimate	95% CrI	BF_{10}	OR	95% CrI	BF_{10}
(Intercept)	−0.02	−3.42–2.87		0.73	0.08–6.98	
Woman candidate	−0.22	−1.08–0.64	3.96	1.28	0.54–3.01	3.71
Negative stimuli	1.31	−1.32–3.85	6.69	1.42	0.22–8.84	2.49
Participant gender	0.89	0.04–1.74	0.54	1.37	0.74–2.57	3.81
Participant ideology	−0.01	−0.45–0.44	0.11	1.00	0.72–1.38	0.08
Participant gender essentialism	0.01	−0.43–0.45	0.12	1.04	0.76–1.43	0.09
Total number of thoughts	0.33	0.17–0.48	100.2	1.11	0.97–1.26	9.43
Woman × negative stimuli	−0.33	−1.72–1.05	2.51			
Negative thoughts				0.63	0.49–0.77	>999
Woman candidate × negative thoughts				1.39	1.09–1.83	275.5
$N_{obs}/N_{id}/N_{stimulus}$	226/70/4			224/70/4		
$ICC/R^2_{marginal}/R^2_{conditional}$	0.26/0.178/0.239			0.14/0.189/0.193		

decision, for example by mentioning women's descriptive underrepresentation or the need for women quota in leadership positions. Participants explicitly use these considerations to rationalize their vote choice, as illustrated by this thought by a male participant (42 years): "If I'm going to have to vote for somebody incompetent, might as well be a woman given there are too few of them."

3. Study 2: Thought-Listing Replication

A frequent criticism of the TA paradigm is that verbalization affects the thinking process, leading to a distorted assessment of cognitive processes (for a review, see Fox et al., 2011). Because Study 1 involved the presence of an experimenter, another concern is that social desirability might drive which thoughts are verbalized. To address these issues, I replicate the design of the first study as a "silent" thought-listing (TL) survey experiment (Erisen et al., 2014; Lodge & Taber, 2013).

3.1. Participants

A total of 142 participants (43% women, M_{age} = 30.7, SD_{age} = 9.6) were recruited via Amazon Mechanical Turk. Participants had to be at least 18 years old and fluent in German. Participation lasted on average 13.3 minutes (*SD* = 7.2) and was rewarded with 1 USD.

3.2. Design and Stimuli

The design and materials were identical to Study 1, except for the following changes. T1 instructed participants to perform the TL task in two steps for each image. First, they viewed images and listed their thoughts as spontaneously as possible in empty text boxes (with a forced list of five thoughts). Second, they then saw their own listed thoughts and classified each thought as either positive, neutral, or negative. For T2, participants moved a slider to either the left or the right to indicate their vote preference for the candidate on the corresponding side (see Panel C in Figure 1).

3.3. Measures

I measure *negativity* as the sum of thoughts that participants classified as negative per candidate image (ranging from 0 to 5, *M* = 1.8, *SD* = 1.6). Participants' *vote preference* is captured on a scale from −50 (preference against candidate) to 50 (preference for candidate), where the scale midpoint of 0 indicates a neutral undecided preference (*M* = 0.7, *SD* = 26.7). The same independent and control variables were used as in Study 1.

3.4. Data Analysis

The unit of analysis was the individual candidate image (*n* = 456). I repeat the same Bayesian multilevel regression models from Study 1, again clustered around the stimulus set and the individual participant.

3.5. Results

The results from Study 2 largely mirror those of Study 1. Model 1 in Table 3 suggests very strong evidence for the absence of an effect of candidate gender on the number of listed negative thoughts (BF_{10} = 0.02). Participants list on average 1.75 more negative thoughts in response to candidate images with negative cues compared to those with neutral cues (CrI = 0.12–2.79, BF_{10} = 12.5) irrespective of candidate gender.

Again, negativity in voter thoughts has a strong negative effect on vote preference, diminishing the preference by 2.90 per listed negative thought (CrI = −4.56–1.23, BF_{10} > 999). Moreover, the interaction effect points to a protective effect where negative

Table 3. Results of Bayesian multilevel regression models predicting negativity and vote preference in Study 2.

	Model 1: Negative Thoughts (T1)			Model 2: Vote Preference (T2)		
Predictors	Estimate	95% CrI	BF_{10}	Estimate	95% CrI	BF_{10}
(Intercept)	1.16	−0.03–2.45		4.23	−4.91–13.38	
Woman candidate	−0.01	−0.34–0.33	0.02	1.09	−2.42–4.62	0.96
Negative stimuli	1.75	0.12–2.79	12.5	0.11	−3.79–4.20	0.92
Participant gender	0.12	−0.22–0.47	0.11	0.08	−3.00–3.16	0.79
Participant ideology	−0.10	−0.24–0.04	0.11	−0.09	−1.59–1.40	0.39
Participant gender essentialism	0.11	−0.03–0.25	0.14	0.18	−1.58–1.96	0.48
Woman × negative stimuli	0.25	−0.22–0.73	0.21			
Negative thoughts				−2.90	−4.56–1.23	>999
Woman candidate × negative thoughts				1.54	−0.14–3.21	27.50
$N_{obs}/N_{id}/N_{stimulus}$	456/82/4			456/82/4		
$ICC/R^2_{marginal}/R^2_{conditional}$	0.22/0.320/0.424			0.81/0.043/0.048		

thoughts are less detrimental to women than men candidates' vote preference (BF_{10} = 27.50). Panel B of Figure 2 shows that women candidates retain a slightly positive vote preference despite the presence of negative thoughts while only little negativity causes a significant drop in preference for men candidates.

4. Overall Discussion and Conclusion

The goal of this study was to investigate the relationship between candidate gender and negativity in voters' evaluation process. I examined voter thoughts in response to neutral and negative candidate cues by means of a mixed methods approach, combining a quantitative and qualitative TA (Study 1) and TL (Study 2) design. Across both studies, a similar pattern emerges: (a) Negative cues elicit the same amount of negativity in voters' thoughts for women and men politicians, (b) these negative thoughts strongly lower candidates' electoral chances, (c) but less so for women candidates.

First, voters' tendency to think negatively of candidates irrespective of gender can be interpreted as an equalizing effect of negativity. One interpretation is that negative cues have primacy over gender cues in the initial, mostly implicit stages of the candidate evaluation process (see, e.g., Lodge & Taber, 2013). This view is supported by psychophysiological studies which have consistently linked negative cues—but not gender cues—to implicit affective responses to political cues (Bakker et al., 2021; Soroka et al., 2019). Moreover, research on affect contagion has shown that initial affective responses spread and favor the activation of similarly charged mental concepts, which are retrieved from memory and made available for further (explicit) processing (Erisen et al., 2014; Lodge & Taber, 2013), including verbalization. I find evidence for this cascading effect of negative cues on further processing both in the quantitative (effect of negative cues on number of negative thoughts) and the qualitative analysis (see block quote in Section 2.4). As negativity selectively reinforces negative thoughts, the activation of gender-related aspects becomes less likely thus reducing their availability as heuristics. However, even if negativity affects the evaluation of women and men candidates similarly, the fact that content analyses have shown more negative media coverage for women politicians (see Van Der Pas & Aaldering, 2020) remains problematic, as this provides more opportunity for negative affect priming (Meffert et al., 2006).

Second and in line with meta-analytic findings (Lau et al., 2007), I find very strong evidence that inducing negativity in voters' thoughts does not win votes. This has implications for candidates' campaign strategies. Though negativity is a losing game for all candidates in this study, the context of actual campaign negativity may modulate how voters think about specific forms of negativity (for a review see Nai, 2020). For example, studies show that voters are less likely to electorally punish candidates who respond to negativity rather than instigating it (Craig & Rippere, 2016; Krupnikov & Bauer, 2014).

Third, the finding of a protective (or less detrimental) effect of negative thoughts for women candidates shifts the mixed empirical evidence on the relationship of negativity and gender ever so slightly towards a more optimistic narrative for women: While detrimental in absolute terms, women suffer less from negative thoughts relative to men (Fridkin et al., 2009; Gordon et al., 2003).

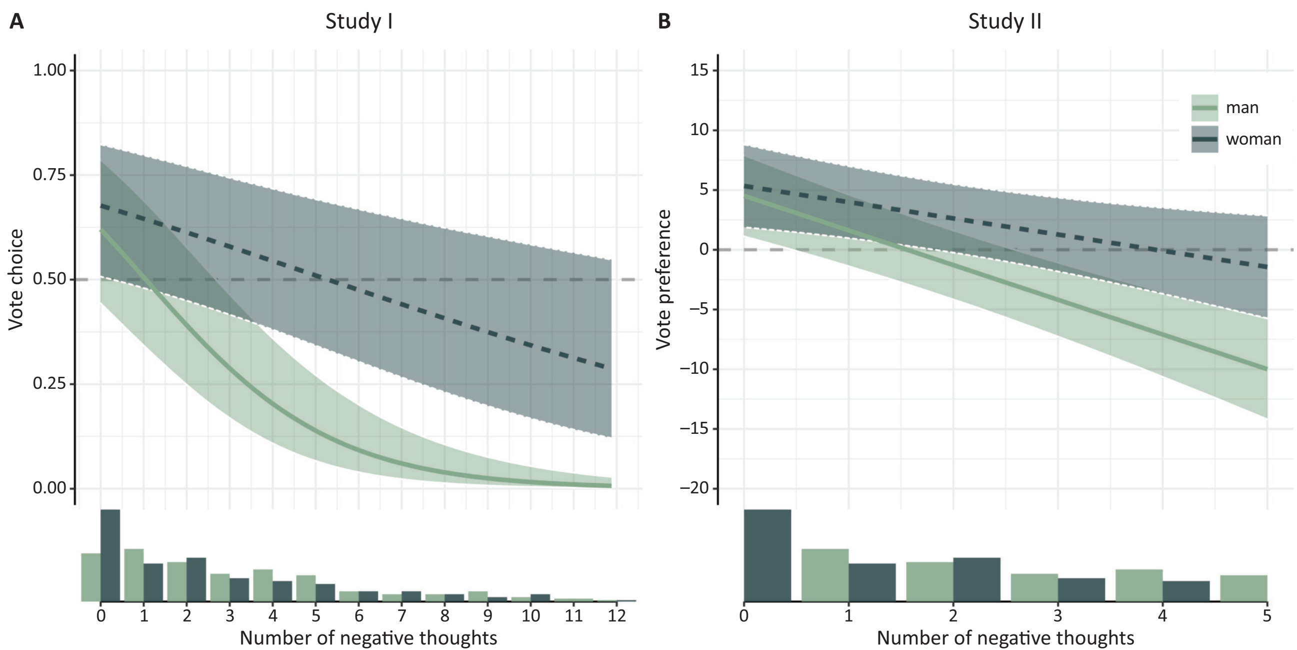

Figure 2. Illustration of the interaction effect of negativity in voter thoughts and candidate gender on vote choice (Panel A; think aloud study) and vote preference (Panel B; thought listing study).

The qualitative insights illustrate that voters frequently invoke women's descriptive underrepresentation when faced with a choice between two candidates and that this perceived power imbalance can tip the scale in women's favor. These explicit gender references at the later stage of the evaluation process (T2) contrast with the earlier impression formation stage (T1) where mentions of gender are scarce. This could mean that negative and gender cues enter the evaluation process at different stages and in different ways. Whereas negativity drives and affectively anchors the initial (implicit) processing of a candidate image, gender marks the context for the (explicit) rationalization of the vote decision. This finding underlines the important role of public perceptions of women in politics for opinion formation (Stauffer, 2021) and adds to recent research suggesting that actively reminding voters of existing biases can be a viable strategy for women candidates (Brooks & Hayes, 2019).

This study comes with several limitations. I focus on explicit dimensions of voter thinking and thus of the candidate evaluation process. This choice implies that any assumptions regarding implicit aspects of candidate evaluation—though established in the literature—remain untested. A promising approach for future studies could lie in the combination of the TA paradigm with implicit approaches, namely psychophysiological measures or implicit association tasks. Moreover, it also raises the issue of social desirability, which could encourage participants to exaggerate their gender perceptions despite methodological efforts to mask the gender-specific goal of the study (through gender-neutral cover stories, filler tasks, and same-gender stimulus sets) or enhancing the anonymity of thoughts (Study 2). However, rationalizations cannot (and should not) be isolated from their social context as they are precisely indicative of how voters reconcile social expectations—such as gender norms—with their own prior attitudes and beliefs (Lodge & Taber, 2013; Yong et al., 2021). Finally, although the design of this study cannot establish (or reject) any underlying mechanism, the protective effect implies that voters are somehow motivated to rationalize away part of the negativity for women but not men candidates. Whether they do so out of benevolent sexism (Barnes et al., 2020; Cassese & Holman, 2019), because they found ways to resolve perceived role incongruence (Bauer, 2017), or following a genuine desire to undo structural inequality remains an open question.

Acknowledgments

For helpful comments on earlier drafts, I thank Katelyn Stauffer, Amanda Friesen, Nandita Biswas Mellaphy, participants and organizers of the workshop on Negative Politics, members of the Hot Politics Lab at the University of Amsterdam, as well as panels at the 2022 Biennial European Conference on Politics and Gender and the 2022 Annual Meeting of the International Society for Political Psychology.

Conflict of Interests

The author declares no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Bakker, B. N., Schumacher, G., & Rooduijn, M. (2021). Hot politics? Affective responses to political rhetoric. *American Political Science Review*, *115*(1), 150–164. https://doi.org/10.1017/S0003055420000519

Barnes, T. D., Beaulieu, E., & Saxton, G. W. (2020). Sex and corruption: How sexism shapes voters' responses to scandal. *Politics, Groups, and Identities*, *8*(1), 103–121. https://doi.org/10.1080/21565503.2018.1441725

Bauer, N. M. (2015). Emotional, sensitive, and unfit for office? Gender stereotype activation and support female candidates. *Political Psychology*, *36*(6), 691–708. https://doi.org/10.1111/pops.12186

Bauer, N. M. (2017). The effects of counterstereotypic gender strategies on candidate evaluations. *Political Psychology*, *38*(2), 279–295. https://doi.org/10.1111/pops.12351

Bridgewater, J., & Nagel, R. U. (2020). Is there cross-national evidence that voters prefer men as party leaders? No. *Electoral Studies*, *67*, Article 102209. https://doi.org/10.1016/j.electstud.2020.102209

Brooks, D. J., & Hayes, D. (2019). How messages about gender bias can both help and hurt women's representation. *American Politics Research*, *47*(3), 601–627. https://doi.org/10.1177/1532673X18795608

Cassese, E. C., & Holman, M. R. (2018). Party and gender stereotypes in campaign attacks. *Political Behavior*, *40*(3), 785–807. https://doi.org/10.1007/s11109-017-9423-7

Cassese, E. C., & Holman, M. R. (2019). Playing the woman card: Ambivalent sexism in the 2016 U.S. presidential race. *Political Psychology*, *40*(1), 55–74. https://doi.org/10.1111/pops.12492

Craig, S. C., & Rippere, P. S. (2016). He said, she said: The impact of candidate gender in negative campaigns. *Politics & Gender*, *12*(2), 391–414. https://doi.org/10.1017/S1743923X15000550

Dolan, K. A. (2014). *When does gender matter? Women candidates and gender stereotypes in American elections*. Oxford University Press.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598. https://doi.org/10.1037//0033-295X.109.3.573

Ericsson, A. K., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* The MIT Press.

Erisen, C., Lodge, M., & Taber, C. S. (2014). Affective contagion in effortful political thinking. *Political Psychology*, *35*(2), 187–206. https://doi.org/10.1111/j.1467-9221.2012.00937.x

Fearon, J. D., & Laitin, D. D. (2013). Integrating qualitative and quantitative methods: Putting it together again. In R. E. Goodin (Ed.), *The Oxford handbook of political science* (pp. 1166–1186). Oxford University Press.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, *137*(2), 316–344. https://doi.org/10.1037/a0021663

Fridkin, K. L., & Kenney, P. J. (2012). The impact of negative campaigning on citizens' actions and attitudes. In M. Scammel & H. A. Semetko (Eds.), *The SAGE handbook of political communication* (pp. 173–185). SAGE.

Fridkin, K. L., Kenney, P. J., & Woodall, G. S. (2009). Bad for men, better for women: The impact of stereotypes during negative campaigns. *Political Behavior*, *31*(1), 53–77. https://doi.org/10.1007/s11109-008-9065-x

Geer, J. G. (2012). The news media and the rise of negativity in presidential campaigns. *PS: Political Science & Politics*, *45*(3), 422–427. https://doi.org/10.1017/S1049096512000492

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC Press.

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. https://doi.org/10.1037/0022-3514.70.3.491

Gordon, A., Shafie, D. M., & Crigler, A. N. (2003). Is negative advertising effective for female candidates? An experiment in voters' uses of gender stereotypes. *Harvard International Journal of Press/Politics*, *8*(3), 35–53. https://doi.org/10.1177/1081180X03008003003

Hoijtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*, *24*(5), 539–556. https://doi.org/0.1037/met0000201

Hughes, M. M., & Paxton, P. (2019). The political representation of women over time. In S. Franceschet, M. L. Krook, & M. Tan (Eds.), *The Palgrave handbook of women's political rights: Gender and politics* (pp. 33–51). Palgrave Macmillan.

Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799. https://doi.org/10.1038/s41593-020-0660-4

King, J. D., & McConnell, J. B. (2003). The effect of negative campaign advertising on vote choice: The mediating influence of gender. *Social Science Quarterly*, *84*(4), 843–857. https://doi.org/10.1046/j.0038-4941.2003.08404012.x

Klein, J. G., & Ahluwalia, R. (2005). Negativity in the evaluation of political candidates. *Journal of Marketing*, *69*(1), 131–142. https://doi.org/10.1509/jmkg.69.1.131.55509

Krupnikov, Y., & Bauer, N. M. (2014). The relationship between campaign negativity, gender and campaign context. *Political Behavior*, *36*(1), 167–188. http://dx.doi.org/10.1007/s11109-013-9221-9

Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. *The Journal of Politics*, *69*(4), 1176–1209. https://doi.org/10.1111/j.1468-2508.2007.00618.x

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge University Press.

Lusk, C. M., & Judd, C. M. (1988). Political expertise and the structural mediators of candidate evaluations. *Journal of Experimental Social Psychology*, *24*(2), 105–126. https://doi.org/10.1016/0022-*1031*(88)90016-9

Maier, J., & Nai, A. (2020). Roaring candidates in the spotlight: Campaign negativity, emotions, and media coverage in 107 national elections. *The International Journal of Press/Politics*, *25*(4), 576–606. https://doi.org/10.1177/1940161220919093

Meffert, M. F., Chung, S., Joiner, A. J., Waks, L., & Garst, J. (2006). The effects of negativity and motivated information processing during a political campaign. *Journal of Communication*, *56*(1), 27–51. https://doi.org/10.1111/j.1460-2466.2006.00003.x

Nai, A. (2020). Going negative, worldwide: Towards a general understanding of determinants and targets of negative campaigning. *Government and Opposition*, *55*(3), 430–455. https://doi.org/10.1017/gov.2018.32

Nai, A., Valli, C., & Maier, J. (2021, May 27–31). *The curse of the iron lady: Reassessing the gender dynamics of negative campaigning via expert, observational and experimental evidence* [Paper presentation]. 71th Annual Conference of the International Communication Association, Virtual Conference.

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, *26*(4), 269–281. https://doi.org/10.1111/1471-6402.t01-1-00066

Schneider, M. C., & Bos, A. L. (2014). Measuring stereotypes of female politicians. *Political Psychology*, *35*(2), 245–266. https://doi.org/doi:10.1111/pops.12040

Schwarz, S., & Coppock, A. (2022). What have we learned about gender from candidate choice experiments? A meta-analysis of 42 factorial experiments. *The Journal of Politics*, *84*(2), 655–668. https://doi.org/0.1086/716290

Sigelman, L., & Kugler, M. (2003). Why is research on the effects of negative campaigning so inconclusive? Understanding Citizens' perceptions of negativity. *The Journal of Politics*, *65*(1), 142–160. https://doi.org/10.1111/1468-2508.t01-1-00007

Soroka, S., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, *116*(38), 18888–18892. https://doi.org/10.1073/pnas.1908369116

Stauffer, K. E. (2021). Public perceptions of women's inclusion and feelings of political efficacy. *American Political Science Review*, *115*(4), 1226–1241. https://doi.org/10.1017/S0003055421000678

Swigger, N., & Meyer, M. (2019). Gender essentialism and responses to candidates' messages. *Political Psychology*, *40*(4), 719–738. https://doi.org/10.1111/pops.12556

Valli, C., & Nai, A. (2020). Attack politics from Albania to Zimbabwe: A large-scale comparative study on the drivers of negative campaigning. *International Political Science Review*. Advance online publication. https://doi.org/10.1177/0192512120946410

Van Der Pas, D. J., & Aaldering, L. (2020). Gender differences in political media coverage: A meta-analysis. *Journal of Communication*, *70*(1), 114–143. https://doi.org/10.1093/joc/jqz046

Van Someren, M., Barnard, Y. F., & Sandberg, J. (1994). *The think aloud method: A practical approach to modelling cognitive*. Academic Press.

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., & Epskamp, S. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Yong, J. C., Li, N. P., & Kanazawa, S. (2021). Not so much rational but rationalizing: Humans evolved as coherence-seeking, fiction-making animals. *American Psychologist*, *76*(5), 781–793. https://doi.org/10.1037/amp0000674

About the Author

Tobias Rohrbach is a joint-PhD candidate at the University of Fribourg and the University of Amsterdam. His research focuses on the intersection of gender and media in political communication, political psychology, and health communication. He specializes in mixed methods design combining a wide range of methodological approaches, including observational and experimental designs, quantitative and qualitative analysis, as well as computational methods.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 384–395
https://doi.org/10.17645/pag.v10i4.5789

Article

Grievance Politics: An Empirical Analysis of Anger Through the Emotional Mechanism of *Ressentiment*

Tereza Capelos [1,*], Mikko Salmela [2,3], and Gabija Krisciunaite [1]

[1] Department of Political Science and International Studies, University of Birmingham, UK
[2] Center for Subjectivity Research, University of Copenhagen, Denmark
[3] Faculty of Social Sciences, University of Helsinki, Finland

* Corresponding author (t.capelos@bham.ac.uk)

Submitted: 13 May 2022 | Accepted: 21 October 2022 | Published: 30 December 2022

Abstract

In this article, we undertake an empirical examination of the psychology of what is often called "the angry citizen," highlighting *ressentiment* as an important emotional mechanism of grievance politics. Contrary to the short-lived, action-prone emotion of anger proper, *ressentiment* transmutes the inputs of grievance politics like deprivation of opportunity, injustice, shame, humiliation, envy, and inefficacious anger, into the anti-social outputs of morally righteous indignation, destructive anger, hatred, and rage. Our empirical probe uses qualitative and quantitative analysis of 164 excerpts from interviews with US "angry citizens" from the following works: *Strangers in Their Own Land: Anger and Mourning on the American Right* (2016) by Arlie Russell Hochschild, *Angry White Men: American Masculinity at the End of an Era* (2017) by Michael Kimmel, and *Stiffed: The Roots of Modern Male Rage* (2019) by Susan Faludi. In these seemingly "angry" excerpts, we find markers matching the psychological footprint of *ressentiment* instead of anger proper: victimhood, envy, powerlessness; the defenses of splitting, projection, and denial; and preference for inaction, anti-preferences, and low efficacy. We conclude on the significance of the distinction between anger proper and *ressentiment* for understanding the psychology of grievance politics.

Keywords

anger; angry citizen; emotional mechanism; grievance; philosophy; political psychology; *ressentiment*; resentment; United States

Issue

This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Contemporary politics is angry and vengeful, with affective polarization and uncompromising antagonisms posing a significant challenge for democracies. We examine the psychology of what is often called "the angry citizen," identifying *ressentiment* as a significant emotional mechanism of grievance politics, distinct from anger proper and aversive affectivity more broadly. Grievance politics refers to the mode of relating to politics primarily through grievances, felt as deprivation of opportunity, injustice, shame, humiliation, envy, and inefficacious anger, and it has recently been the focus of a growing number of studies. Salmela and von Scheve (2018) elaborate on the pro-social forms of grievance politics discussing civil rights and LGBTQ social movement dynamics through the emotional mechanism of *social sharing*. Capelos et al. (2021) examine the relationship between the backward gaze of reactionism, its *ressentiment*ful affective core, and collective narcissism with its

precarious social bonds, labeling them "the anti-social triad of grievance politics." These studies make an important distinction between prosocial and antisocial forms of political grievance. Salmela and Capelos (2021) theorize *ressentiment* as the emotional mechanism that transmutes the inputs of grievance politics into anti-social outputs of morally righteous indignation, destructive anger, hatred, and rage. Flinders and Hinterleitner (2022) discuss the decline of party politics and the rise of grievance politics. Capelos and Demertzis (2018, 2022) examine the dormant support for violent political action among *ressentiment*ful citizens, their hollow social contact, precarious collective identities, and their negative relationship with political knowledge, scientific evidence, and emancipatory values, joining recent studies which argue the central role of *ressentiment* in contemporary far-right, populist, and nationalist contexts (Ciulla, 2020; Demertzis, 2020; Pirc, 2018; TenHouten, 2018; Wimberly, 2018).

Ressentiment is not a new concept. Originating from Nietzsche (1885/1961) and elaborated by Scheler (1915/1961), it is applied in studies of extremism and fundamentalism (Posłuszna & Posłuszny, 2015; Žižek, 2008), Trumpism (Knauft, 2018; Wimberly, 2018), fanaticism (Katsafanas, 2022), right-wing populism (Salmela & von Scheve, 2017, 2018), reactionism (Capelos & Demertzis, 2018; Capelos & Katsanidou, 2018; Sullivan, 2021), narcissism (Demertzis, 2020), terrorism (Posłuszna, 2019, 2020), extremism (Mishra, 2017), and cynicism (Capelos et al., 2021; Halsall, 2005). Drawing from studies of emotional mechanisms and their key function of transforming an input emotion into a different output emotion (Elster, 1999; Salice & Salmela, 2022), Salmela and Capelos (2021) approach *ressentiment* as an *emotional mechanism* that transmutes political, social, or private grievances felt as deprivation of opportunity, injustice, humiliation, and lack of political efficacy, to anti-social emotional expressions of morally righteous indignation, destructive anger, hatred, and rage.

We use this conceptualization of *ressentiment* to elaborate on the psychological nature of contemporary "angry politics." We distinguish the *inputs* of grievance politics, such as political disaffection, frustration, deprivation, and discontent, from the *outputs* of grievance politics which can be (a) constructive outputs with collective action potential, delivered through social sharing and (b) anti-social, maladaptive, bitter, and vengeful outputs delivered via *ressentiment*. Angry politics founded on grievance can impart pro-social social change (Salmela & von Scheve, 2018), whereas *ressentiment*ful politics founded on grievance are antisocial (Salmela & Capelos, 2021). Fundamentally, we argue, the problem in the study of grievance politics is how to tell apart anger proper from *ressentiment*.

We approach this challenge in three steps. First, we distinguish between anger proper and the vengeful, inefficacious venting of frustrations towards out-groups denoting *ressentiment*. Second, we engage with the deep psychological processes of *ressentiment* and discuss its employment of defenses, the unconscious mental processes which, through emotional self-adjustment, serve as an invisible "defensive shield" from intrapsychic conflicts, and their affects (Cramer, 2015; Vaillant, 1993). Third, we examine the expressions of *ressentiment* among individuals widely perceived as *angry*. We analyze the content of displays of anger and *ressentiment* in 164 excerpts of interviews with US "angry" citizens sourced from *Angry White Men* (Kimmel, 2017), *Strangers in Their Own Land* (Hochschild, 2016), and *Stiffed: The Roots of Modern Male Rage* (Faludi, 2019). We find key constitutive markers of *ressentiment* (envy, victimhood, powerlessness, destiny, transvaluation, and injustice) in expressions broadly understood as anger, and evidence of its inefficacious approach to politics delivered through inaction and anti-preferences.

The important differences between anger and *ressentiment* elucidate the puzzle of bitter disengagement and alienation from democratic representation which has become a hallmark of contemporary politics. Our article contributes to studies focusing on emotions to understand negativity and affective polarization in politics (Brader, 2006; Gadarian & Albertson, 2012; Huddy et al., 2002, 2008; Jost et al., 2017; Mason et al., 2021; Mayer & Nguyen, 2021; Turner, 2007), the rise of authoritarian and far-right populist leaders, and the processes by which animus and antagonistic politics gain ground in post-truth electoral campaigns (Achen & Bartels, 2016; Betz, 1993; Farkas & Schou, 2019; Forgas et al., 2021; Kisić Merino et al., 2021; Marcus, 2021; Michel et al., 2020; Nai, 2021; Norris & Inglehart, 2019). We also add to the growing number of empirical studies on *ressentiment* which have examined political experiences in populist contexts (Capelos & Demertzis, 2018, 2022; Capelos et al., 2021; Ciulla, 2020; Demertzis, 2020; Kazlauskaitė & Salmela, 2021; Mishra, 2017).

2. Anger vs. *Ressentiment*: The Complexities of Repressed Aggression

Anger and *ressentiment* are conceptually and psychologically close but they are not the same emotional experience (Meltzer & Musolf, 2002; Solomon, 1995). Anger is a discrete emotion with a defined object, usually short-lived, and generates action tendencies. We refer to it herein as "anger proper" to distinguish it from general accounts of negative emotionality. Anger-proper arises as a response to the appraisal of an event which is not in the individual's control, seen as an obstruction or infringement to reaching a goal or satisfying a need. It is bound to personal or social expectations, and results in physiological changes and mental readiness, which prepare an individual for action (Capelos, 2013; Ekman, 2004; Frijda, 2004; Lazarus, 1993; Roseman & Evdokas, 2004). It is associated with negative reactive attitudes towards political objects, decline in political trust, weakened commitment to democratic norms and values,

optimistic risk estimates, out-group hostility and racial aggression, increased discontent, and desire to punish (Brader et al., 2008; Gadarian & Albertson, 2012; Huddy et al., 2008; Lerner et al., 2003; Phoenix, 2019; Webster, 2020). Many contemporary philosophers of emotion consider anger a healthy and appropriate response to unjust or unfair circumstances (Huddleston, 2021; Thompson, 2006; yet for a critique of anger, see Nussbaum, 2016).

If one must think of *ressentiment* in terms of anger, then the closest approximations are inefficacious anger and blunted vindictiveness of a toxic kind. Yet again, *ressentiment* is more complex. According to Nietzsche (1885/1961) and Scheler (1915/1961), *ressentiment* is a largely unconscious experience which works primarily as a "psychic shield" from negative emotions and feelings of injustice and humiliation, as well as deprivation from the desired, with a shadow of inferiority. Salmela and Capelos (2021) offer a consolidated review of theories of *ressentiment*, and define it as a long-lasting compensatory emotional mechanism, triggered by envy, shame, or inefficacious anger, all involving a sense of self-reproaching victimhood. Unlike the short-lived character of anger, *ressentiment* has a lasting impact on the individual, as it involves a transvaluation of the self and its values. It is inefficacious and vengeful, it employs defenses, and it is dynamic: It transmutes lacerating emotions like envy, shame, and inefficacious anger into outcome emotions of moral anger (as resentment, when felt about personal wrongs towards one's self or one's people with the desire for personal revenge; or as indignation, when felt about impersonal wrongs with the desire to see wrongs righted by a third-party punishment; see Aeschbach, 2017, pp. 30–37) and hatred, displaying a morally superior victim position.

While resentment as moral anger can emerge on its own or through *ressentiment*, it is important to observe differences in the intentional targets and action tendencies between these two types of resentment. The first type of resentment is moral anger at injustices and wrongs that motivates individual or collective action seeking to correct or retribute the relevant injustice or wrongdoing. This high action readiness associates the first type of resentment with anger-proper. The second type of resentment resulting from *ressentiment* is more complex as it is generated from repressed shame, envy, or humiliation, which are intolerable for the self (Lewis, 1971; Scheff,1994; Scheler, 1915/1961; Turner, 2007). Therefore, resentment mediated by *ressentiment* has an indeterminate and "blurred" affective focus on generic "enemies" of the self (cf. Szanto, 2018) that allows its targeting to various scapegoats in political rhetoric (Salmela & von Scheve, 2017). Furthermore, the resentment felt via *ressentiment* is not an active emotion of protest associated with anger-proper (Jasper, 2014; Salmela & von Scheve, 2018). It is inefficacious, closer to revenge "taken on the object in thought rather than in action" (Nietzsche in Hoggett, 2018, p. 394; Scheler, 1915/1961). In their collective political expressions, the transition from *ressentiment* to anger-proper can be transformative. As Kiss (2021) notes, when political leaders and institutions function as discharge mechanisms, grievance politics of *ressentiment* can be transformed into anger, changing society from passive to active.

Theoretical accounts of *ressentiment* highlight the role of envy, shame, and inefficacious anger as its trigger emotions. In addition, they recognize the central role of victimhood, powerlessness, the process of transvaluation, and a strong sense of destiny (Aeschbach, 2017; Salmela & Capelos, 2021). Phenomenological and macro-historical sociological approaches (Demertzis, 2006; Ferro, 2010; Moruno, 2013; Scheler, 1915/1961; Szanto & Slaby, 2020) and empirical studies on *ressentiment* focus on victimhood, envy, powerlessness, destiny, and transvaluation as its key markers to distinguish it from anger proper. León et al. (1988) created a 28-item survey scale with items measuring envy, victimhood, indignation, powerlessness, sense of injustice, and destiny. Capelos and Demertzis (2022) used a shortened six-item version of this scale, while Capelos and Demertzis (2018) relied on a proxy measure which combined anger, anxiety, and low political efficacy.

One shortcoming of extant measures of *ressentiment* is that they are static, while emotional mechanisms are dynamic. Salmela and Capelos (2021) proposed the empirical measurement of *ressentiment* via the observation of defenses alongside its key markers and highlighted the value of the defenses of *reaction formation* (as the exaggerated opposition and preoccupation with the object of desire), *splitting* (the world is all good/the world is all bad; I am all good/the other is all bad), *denial of facts* (refusal to accept reality), *projection* (bad elements of the self are projected out), *introjection* (good elements of external objects are incorporated with the self), and *mirroring/idealization* (strong identification with the other as a mirror to oneself). While defenses operate at the level of individual psychology, they are supported and reinforced by public discourses in traditional and social media by opinion leaders and political entrepreneurs whose affective rhetoric contributes to the transvaluation process (Kazlauskaitė & Salmela, 2021). The use of defenses as proxies of *ressentiment* can therefore apply to individual and group level studies.

Ressentiment can also be identified through its outcome emotions. As the "new-self" with its "new values" seeks validation through social sharing with peer-others (consolidating stage) the hostile outcome emotions, hatred and moral anger (resentment, indignation), typically directed towards concrete objects, such as other persons or groups, transform (through their lasting reliving and repression) into an objectless hostile sentiment, easily re-attached to any target, from immigrants and religious groups to government, leaders, elites, or political parties (Aeschbach, 2017; Leiter, 2014; Salmela & Capelos, 2021; Salmela & von Scheve, 2017). This is where *ressentiment* meets negative affectivity (Capelos & Demertzis, 2018) and can be conflated with anger

proper. They have however two important differences: The generalized toxic target-emotionality of *ressentiment* is perceived as morally *righteous* (Salmela & Capelos, 2021), and it is shared with one's peers as an "antagonistic affective attachment" between the individual and the target of *ressentiment* (Szanto & Slaby, 2020, p. 15).

The expression of *ressentiment* in grievance politics goes beyond anger proper, and is linked to victimhood, powerlessness, inefficacy, and inaction (Capelos et al., 2021; Salmela & Capelos, 2021). Individuals in *ressentiment* display morally righteous indignation which gives rise to "victimological collectives" (Sloterdijk, 2010, p. 152) but will not actively engage in democratic participatory acts or collective actions (Capelos & Demertzis, 2018; Hoggett, 2018; Salmela & Capelos, 2021). As if the present is not worth engaging in, and the future is distantly disconnected from their grievances, individuals in *ressentiment* remain attached to nostalgic accounts of the past. Their bitterness is expressed as dogmatic, binary anti-preferences, sustained through lasting rumination over remembered or imagined injustices (Capelos & Demertzis, 2018). When populist and authoritarian leaders co-opt bitter and retrogressive/nostalgic narratives and agendas, *ressentiment* becomes politically relevant (Capelos et al., 2021). Studying the psychology of *ressentiment* allows us to recognize how past or present perceived injustices are gradually internalized by individuals or collectives as if these were a constituent part of their identity making them special (Adler, 2013). In *ressentiment*, nostalgia does not simply denote "a longing for a home that no longer exists" (Boym, 2001, as cited in Reynolds, 2004, p. 2). *Ressentiment*ful nostalgia is bitter. It manifests as grievance projected on out-groups and results in anti-stances including anti-feminist, anti-immigration, anti-globalization, anti-science, anti-elite, and anti-EU positions (Capelos & Katsanidou, 2018; Capelos et al., 2017; Ciulla, 2020; Sullivan, 2021).

The nuanced psychological composition and behavioral expressions of anger proper and *ressentiment* justify their conceptual and empirical differentiation. While telling them apart is not an easy task, we rely on strong theoretical insights to generate distinct markers for each concept: (a) Anger proper and *ressentiment* can share a sense of injustice, but, contrary to anger proper, *ressentiment* is inefficacious and passive and is tapped through markers not applicable to anger: moral victimhood, envy, powerlessness, destiny, and transvaluation; (b) *ressentiment* involves defenses of projection, introjection, splitting, denial, and reaction formation, whereas anger has no theorized relationship with defenses; (c) *ressentiment* is marked by inefficacy, inaction, and anti-preferences. Anger proper, occurring outside the emotional mechanism of *ressentiment*, would not display these markers to the same degree and as consistently.

We expect that what is often perceived as "anger" against political elites, the establishment, and "enemy others," will have a *ressentiment*ful core. We employ an empirical plausibility probe to establish the validity of this expectation. A plausibility probe is a stage of empirical inquiry preliminary to testing, which examines the plausibility of a theory. Empirical plausibility probes adopt suggestive tests, do not require large representative samples, and establish whether a theoretical construct is worth considering, without providing exact estimates of probability (Eckstein, 1992).

3. Empirical Plausibility Probe

Our empirical plausibility probe involves qualitative and quantitative content analysis (Burla et al., 2008; Green et al., 2007; Rourke & Anderson, 2004; Schreier, 2014; Vaismoradi et al., 2016) of 164 excerpts from interviews with "angry" citizens in the US. We sourced interview excerpts from three books focusing on contemporary expressions of anger in politics: *Angry White Men: American Masculinity at the End of an Era* by Kimmel (2017), *Strangers in Their Own Land* by Hochschild (2016), and *Stiffed: The Roots of Modern Male Rage* by Faludi (2019). Kimmel (2017, p. 9) discusses "the rage of the American 'every-man"' focusing on middle-class men and the sources as well as objects of their anger. Hochschild (2016) investigates the right-wing supporters' stance towards American politics in Louisiana. Faludi (2019) discusses how rising challenges to the traditional understanding of masculinity in the late 1990s led not only to gender wars but also to accumulated anger which can still be felt in contemporary America.

The books contain interviews collected in the US and address anger in relation to issues in the private, public, and political realms with middle-class citizens, Tea Party activists, Trump's supporters, white supremacists, fathers' rights activists, Promise Keepers, and others. The authors identify negative affect in the interviews as evidence of anger. Because *ressentiment* was not on the analytical radar of the authors, they did not address it or probe for it in the interviews, making the analysis of the interview excerpts a hard empirical case for the identification of *ressentiment*.

We expected evidence of anger proper and coded for consistent and inconsistent markers. Kimmel (2017, p. 38) briefly refers to *ressentiment* as an emotion of "creative hatred" but does not theorize further. Hochschild (2016, pp. 115, 135, 147, 212) references resentment (not *ressentiment*), specifically the impact of class conflict as the source of resentment in "the American right," considering resentment alongside other emotions (fear, pride, shame, hope, anxiety) to elaborate on the affective experiences of Tea Party activists. The emotional mechanism of *ressentiment* is not explored, but what Hochschild discusses as resentment shares the key markers of *ressentiment*.

Our unit of analysis is each of the 164 excerpts of interviews offered in these books as examples of anger. From all cited interview material, we selected the statements with explicit and implicit mentions of negative

affect. If the text did not contain references to negative affect, we omitted it. Of the 164 negative affect statements (NAS), 71 were drawn from *Angry White Men*, 67 were from *Stiffed*, and 26 were from *Strangers in Their Own Land* (details of the NAS case extraction in Table A1 in Appendix A of the Supplementary File). The number of words varied between 7 and 219 (mean 52). The 164 NAS were made by 108 individuals (1–7 NAS/individual). Most individuals made one statement (86%, 82 of 108). The gender breakdown was 94 men (87%), 12 women (11%), and two (2%) whose gender was not identified. The gender imbalance in our sample is linked to the themes of two of the three books (frustrated masculinities). The topic itself provides a fertile ground to study expressions of anger. It is therefore not surprising it was chosen by the authors, nor is it surprising that the majority of interview subjects were men. We discuss the implications of this in our conclusion.

3.1. Operationalization and Coding

We coded each NAS for instances of anger proper, efficacy, and support for political action, markers of anger in the literature, and non-anger-related measures of *ressentiment*, defenses, inefficacy, inaction, and anti-preferences (coding examples in Tables B1–B4 in Appendix B of the Supplementary File). We applied qualitative content analysis to determine the explicit and implicit meaning of selected texts (Schreier, 2014). To eliminate potential coder bias, intercoder reliability was established across three independent coders on a sample of 10 statements based on 34 key variables. Once satisfactory intercoder reliability was reached (90% agreement across coders), the remaining statements were coded by one coder (Burla et al., 2008).

Anger proper was identified through discrete words like "anger," "rage," and "enraged," and expletives indicating angry frustration, like "screwed." *Ressentiment* was identified by six items adapted from Capelos and Demertzis (2022): *envy* (others do better with less effort), *victimhood* (others take advantage of me), *transvaluation* (reversal of value, from important to unimportant, good to bad), *injustice* (what is happening to me is unfair), *powerlessness* (I feel disrespected), and *destiny* (my hopes will never come true). Except for injustice, which can be an element present in anger, these markers map *ressentiment* and are not consistent with the psychological experience of anger proper.

We coded defenses of *projection* (what is considered bad in the self is projected outwards to another), *introjection* (what is good in the outside world is introjected in the self), *splitting* (oversimplifying reality by splitting the world in all-good and all-bad objects), *denial of facts* (a negation of painful reality), and *reaction formation* (repression of the original affect/desire with the exaggeration of the opposite, like "I am not sad, I am elated"). These defenses are not markers of anger proper but are prominent in *ressentiment*. Coding for defenses is notoriously difficult and scale inventories and deep psychoanalytic techniques acknowledge measurement validity and reliability issues (Soroko, 2014). We consulted validated defenses inventories and studies that identify defenses in interview and narration material (see Hentschel et al., 1993). We recognize that our coding approximates rather than clinically measures these primarily unconscious psychological strategies.

Action tendency (expected high for anger, low for *ressentiment*), was measured as support for action, dormant action (I would support this), and inaction. Efficacy (high for anger, low for *ressentiment*) was measured as a dichotomous yes/no variable of whether individuals stated they were able to influence the event they talked about. Action type recorded whether actions were legal or outside mainstream politics (illegal/violent). To tap into grievance politics, we coded for mentions of anti-preferences (anti-feminist, anti-government, anti-immigration, anti-democracy) and nostalgic thinking.

4. Analysis

To examine the prevalence of anger vs. *ressentiment*, we compared the frequency of anger and *ressentiment* markers across all NAS and between a smaller sample of High*R* (16 NAS containing four or more *ressentiment* markers), and No*R* (28 NAS with no *ressentiment* markers). To understand whether grievance politics was discussed in NAS through anger or *ressentiment*, we looked at the frequencies of inefficacy, inaction, nostalgia, and anti-preferences in the High*R* and No*R* NAS. Because our empirical framework was designed for the secondary analysis of interview excerpts, we provide quantitative tallies of markers of anger proper vs. *ressentiment* and make modest use of the excerpts in the text. We are not able to make extensive use of qualitative content, as we would have in the case of original interview material because we could not probe deeper into the original interviewee's answers. To highlight the rich content of the excerpts and the value for further research, we complement the analyses with relevant excerpts in Tables C1–C5 in Appendix C of the Supplementary File.

4.1. In the Deep: The Emotional Mechanism of Ressentiment and Its Defenses

Across the 164 NAS, only 28 (17%) did not mention victimhood, envy, transvaluation, injustice, powerlessness, or destiny, and 136 (83%) contained one or more of these *ressentiment* markers. Instead of finding mainly language consistent with angry proper in NAS, we counted in total 313 *ressentiment* references which empirically seemed out of place (except injustice) if these individuals were just angry. NAS often contained combinations of *ressentiment* markers: 46 statements (28%) had two, 40 (24%) combined three, 13 (8%) combined four, and three (2%) combined five markers, while 34 (21%) had one marker.

Turning to anger proper, it was present in NAS but less frequently than we expected (84 anger mentions in 164 statements). We also identified only 10 NAS (6%) where anger was present without *ressentiment* (No*R*), and 18 NAS (11%) with negative affect which was neither anger nor *ressentiment* ("If I know a person is a Christian, I know we have a lot in common. I'm more likely to trust that he or she is a moral person than I would a non-Christian,'' NAS 95, pointing to *distrust*).

Injustice, the common link of anger proper and *ressentiment*, was in 12 NAS (4%). The other five markers, properties of *ressentiment* but not of anger proper, were more frequent: Victimhood appeared 116 times (37%), followed by envy (70; 22%), powerlessness (60; 19%), and destiny (42; 13%). Transvaluation was less frequent (13 NAS, 4%) because despite being an important process in *ressentiment*, it is difficult to detect with single-time measures. Understood as the change of value labels over time, transvaluation is often non-conscious and a bad candidate for self-report data. Taken together, the frequent mentions of victimhood, envy, powerlessness, and destiny pointed to a high volume of *ressentiment*, contradicting the expectation that NAS mainly expressed anger proper.

Defenses are expected in *ressentiment* but not in anger proper. We identified 45 NAS (27%) containing a total of 55 mentions of defenses. The most frequently mentioned were projection (19 mentions; 35%) and splitting (17; 31%). Denial (eight; 15%), introjection (six; 11%), and reaction formation (five; 9%) were less frequent. It is logical to anticipate higher frequency of defenses in high *ressentiment* NAS compared to low *ressentiment* NAS. The 16 High*R* NAS contained more frequent splitting (38%) and projection (13%) compared to the 28 No*R* NAS, which showed no splitting (0%), and projection in 11%.

To further examine the theorized link between *ressentiment* and defenses, we examined the *ressentiment* markers present in the defense-containing NAS. We expected core *ressentiment* markers (victimhood, envy, powerlessness) to appear more frequently than injustice which is shared with anger proper. The solid bars in Figure 1 show the totals of *ressentiment* markers across defenses. We see splitting (in orange) and projection (in blue) containing the most *ressentiment* markers, 51 and 45 respectively, confirming the link between *ressentiment* and defenses. The first bar cluster in Figure 1 (in black) shows how inexplicably linked are defenses with the *ressentiment* markers, particularly victimhood (horizontal stripes bar), envy (diagonal upward stripes bar), and powerlessness (diagonal downward stripes). Characteristically, victimhood appeared 47 times in NAS with defenses, envy 35 times, and powerlessness 31 times. These flagship *ressentiment* markers were the most prominently featured compared to the sparse mention of injustice (small grid bar) and transvaluation (dotted bar) in NAS-containing defenses. As injustice is a shared marker of anger, we did not expect to see it frequently with defenses. Transvaluation was very rarely identified in our data overall (see Table D1 in the Supplementary File).

4.2. The Muted and Bitter Voice of Ressentiment: Inefficacy, Inaction, Anti-Preferences, and Nostalgia

Because *ressentiment* originates from inefficacious anger, we expected inefficacy in the 16 High*R* NAS. Indeed, 12 (75%) contained inefficacy alongside mentions of destiny, victimhood, powerlessness, and envy. For example, NAS10 notes: "The inmates are running the asylum. They're completely in power, and they get anything they want. And us regular, normal white guys—We're like nothing. We don't count for shit anymore." NAS68 also expressed inefficacy in the words: "It's like we're nothing….No one listens to the little guy." Characteristically none of the High*R* NAS mentioned efficacy. However, among the 28 No*R* statements, 10 (36%) mentioned efficacy and there were no mentions of inefficacy. An example was NAS125 which reads: "I want to get control of the world. Well, not the world, but I want to get where they see me because I'm on top, where all heads turn when they say my name."

Ressentiment is passive (Capelos & Demertzis, 2018) and as we expected references to political action were sparse. About 75% of High*R* NAS denoted inaction, compared to 61% of No*R* NAS. Instead, action was more frequent in No*R* NAS (25%) compared to High*R* NAS (19%), and was often associated with injustice. This is not surprising given that action is a tendency associated with anger proper which is often a response to a perceived injustice. For example, NAS109 shows action by noting: "All we black union men went to crying because we knew what was going to happen, except for me. When I heard they intended to fire me, I quit before they could."

What we found most interesting was that the inefficacious anger of *ressentiment* was delivered through support for dormant action, which appeared in statements alongside envy and victimhood. The combination of these markers comes through clearly in excerpts like NAS69:

> It is our RACE we must preserve, not just one class….White power means a permanent end to unemployment because, with the non-whites gone, the labor market will no longer be over-crowded with unproductive niggers, spics and other racial low-life. It means an end to inflation eating up a man's paycheck faster than he can raise it because the economy will not be run by OUR criminal pack of international Jewish bankers, bent on using the white worker's tax money in selfish and even destructive schemes.

Anti-preferences can signal political frustration expressed as anger, and grievance expressed through *ressentiment*. Across the 164 NAS, 82 contained anti-preferences, targeting women, immigrants (e.g., black, Hispanic, and Muslim), and the government.

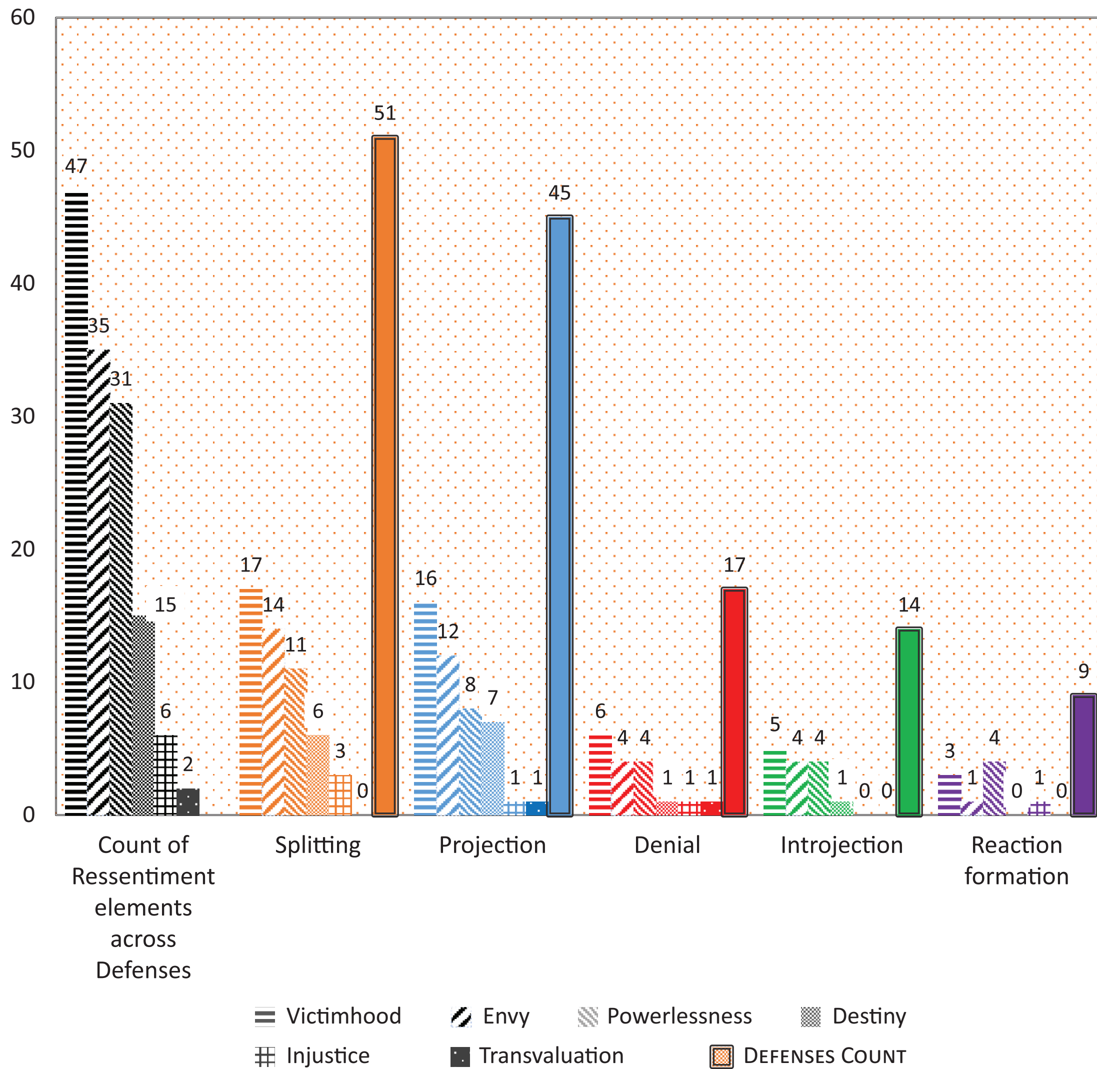

Figure 1. *Ressentiment* and defenses. Notes: Clustered bars show crosstabulations of *ressentiment* markers by defense; the colored bars show counts of ressentiment markets; the first bar cluster (in black) shows counts across defenses, the second bar cluster (in orange) shows counts in splitting, the third bar cluster (in blue) shows counts in projection, the fourth bar cluster (in red) shows counts in denial, the firth bar cluster (in green) shows counts in introjection, the sixth bar cluster (in purple) shows counts in reaction formation; patterned bars show the markers of *ressentiment* and solid bars show total counts of ressentiment markers per defense.

Anti-preferences were more frequent in High*R* (69%) compared to No*R* (46%) NAS. An example of a HighR NAS with anti-immigration, anti-elite, and anti-black references appears in NAS1:

> I mean, just look around. There's illegals everywhere. There's Wall Street screwing everybody. And now there's a goddamn…. Oh, fuck it, I don't care if it is politically incorrect. We got a fucking nigger in the White House.

In NAS21, the attack is against feminists:

> The misandric zeitgeist, the system of feminist governance that most are still loath to acknowledge is about to head toward its inevitable and ugly conclusion, and the results of that will inflict another deep wound on the psyche of the western world.

We also identified 15 references to nostalgia (9%), featuring destiny and sadness for what appears lost. In NAS5, nostalgia is evident in the words: "I liked it the way it was….It's not going to be like that anymore." Another example of nostalgia in NAS57 reads: "Back in the day, if you got screwed by your company, you could go to the government, get unemployment, get food stamps, whatever, get some help. Now there's nowhere to go." One more example of longing for the past in NAS72 reads: "When I was a kid, you stuck a thumb out by the side of the road, you got a ride. Or if you had a car, you gave a ride. If someone was hungry, you fed him. You had community." These seemingly happy memories of

the abandoned heartland are bitter-sweet, laced with grievance and frustration. Unlike the "Make America Great Again" uplifting accounts of populist rhetoric, these *ressentiment*ful recountings imply a lost past which does not come with restoration and is unlikely to return (Sullivan, 2021).

5. Discussion

After finding 83% of NAS containing *ressentiment* markers which were inconsistent with anger proper, we examined the excerpts' psychological content. The prevalence of victimhood came as no surprise, considering the harsh realities of those who talked about losing their jobs, being divorced, being "forced" to pay alimony or feeling betrayed by the government. The content of their statements was determined largely by the context of their life experiences. Expert accounts of *ressentiment* suggest individuals experience "the sacralization of victimhood" (Demertzis, 2017, p. 12), and can get "stuck" in their victimhood status, rather than striving to remove the underlying injustice. In these excerpts, powerlessness was blended with victimhood and the feeling of injustice as destiny, which precludes anything from being done (Demertzis, 2020; Hoggett, 2018; Salmela & Capelos, 2021).

Recognizing the important problem of mistaking *ressentiment* for anger proper, we mapped the defenses it employs, thereby also expanding its instruments. We identified examples of splitting the world into "all-good vs. all-bad": For example, the goodness of one's in-group was defined in contrast to an all-bad out-group. Projection (NAS120, "Girls have all the power"), when coupled with *ressentiment*, focused on victimhood, in line with the moral expression of *righteous* victimhood in *ressentiment* (Hoggett, 2018). Introjection (NAS42, "The knowledge accumulated by men in the ages") also focused on victimhood, which shows that introjection works in feedback loops with projection to reinforce victimhood perceptions (Salmela & Capelos, 2021).

Turning to the political implications of negative affectivity, we examined the theorized link between *ressentiment* and anti-preferences. As expected, the object of *ressentiment* was generalized (Aeschbach, 2017; Salmela & Capelos, 2021), displaced onto one or more "enemy-other(s)," and nostalgia was mostly a hopeless gaze to the past (NAS87, "My grandfather homesteaded those 40 acres before anybody even knew what a refinery was…. It's all killed now. It makes me not want to live in Bayou d'Inde and makes me sad"). Our analysis of efficacy yielded results consistent with theoretical accounts: High*R* NAS, particularly those featuring destiny, mentioned inefficacy and inaction, aligning with Capelos and Demertzis' (2018) findings of *ressentiment* being inefficacious and passive.

Reflecting on our empirical framework, using excerpts of secondary interview material had benefits and drawbacks. We find value in our approach, as it allows the identification of key concepts while avoiding researcher biases related to collecting primary material. Our method is in this sense closer to document and text analysis rather than primary interviews. This comes with limitations: The books we sourced focused primarily on anger and did not aim to uncover *ressentiment*. As a result, the excerpts were not as rich as they might have been if interviews were dedicated to the exploration of *ressentiment* (Hox & Boeije, 2005; Salmela & Capelos, 2021).

Our analysis confirmed how notoriously difficult it is to capture transvaluation with static data (Demertzis, 2020; Hoggett, 2018; Salmela & Capelos, 2021). The muted transformation of one's values and one's sense of self would be more easily discernible through longitudinal data, recording over-time shifts of the values of the self and the objects of *ressentiment*, or through in-depth analysis of qualitative interviews and focus group material where participants elaborate on value changes.

Quantitative measures allow the systematic and parsimonious study of complex phenomena through a relatively small number of indicators. Here we attempted to capture the complex psychological footprint of *ressentiment* with markers tapping on its core drivers, the defenses it employs, and its outcome emotions and experiences. An important assumption when doing this work is that the transformation of emotions and values in *ressentiment* remains unfinished, and therefore driver and outcome emotions are perceivable in the expressions of persons in *ressentiment* (Demertzis, 2020; Salmela & Capelos, 2021). A natural extension is to apply our coding frame to primary interviews and focus group material. This would move the unit of analysis from statements to individuals and groups, opening opportunities for follow-up questions, and the study of non-verbal cues, providing deep meaning through the observation of silences, facial expressions, and body language (Ekman, 2004).

The excerpts we analyzed reflected the experiences of individuals living in the US, the majority of whom were men. Of course, this sample cannot capture the complexity of *ressentiment* in the West, let alone Eastern European states, or states of the Global South. Our findings point to the socially and politically established link between masculinity and anger: Men traditionally express their anger outward and discuss topics they find frustrating, whereas women suppress it, direct it inwards, or sublimate it (Thomas, 2003). As the examination of gender-based and minority differences and similarities is gaining scholarly attention (Kisić Merino et al., 2021; Negra & Leyda, 2021; Phoenix, 2019), collecting geographically and historically diverse material, sampling women and minorities, would advance this line of work.

6. Conclusion

Our take-home point is that anger-focused interpretations of societal dissatisfaction and political grievance can often conceal *ressentiment*, particularly when

individuals feel entitled, yet excluded from a way of life, job, or privileges. This distinction between anger proper and *ressentiment* has significant implications for political life. In *ressentiment*, individuals lack the agency of their angry counterparts. Their inability to publicly express and/or act on their inefficacious anger, envy, or shame, can be very painful and sets forth defenses, fostering rumination and political inaction. Scholars of *ressentiment* are familiar with the original Latin meaning of the term "*re-sentire*,'' to "re-feel" time after time (Hoggett, 2018, p. 395). The other-targeting negative emotions like hatred, resentment, and hostility generated by *ressentiment* promote polarized political preferences and noxious behaviors (Capelos & Demertzis, 2018, 2022).

Studies agree on the *inputs* of grievance politics: The crisis-laced rhetoric of populist parties and the emotional experiences of voters supporting them are rife with aversive affectivity expressed as anger, discontent, pessimism, insecurity, anxiety, blame, and distrust (Betz, 1993; Mudde & Rovira Kaltwasser, 2018; Rico et al., 2017). Elaborating on the psychology of anger vs. *ressentiment* invites the study of emotional mechanisms shaping the political outputs of grievance politics across Western and non-Western populist, nationalist, and authoritarian contexts (Kisić Merino et al., 2021; Sharafutdinova, 2020). The appeal of such rhetoric and narratives, particularly on the far-right, feeds and grows through subjective and intersubjective perceptions of threat and vulnerability (Kinnvall & Svensson, 2022; Salmela & von Scheve, 2017). Crucially, the outputs of grievance in *ressentiment* are not the outputs of grievance in anger. The long-lasting anti-social, vengeful, moral victimhood of *ressentiment* is distinct from the collective action potential of anger generated by social sharing of frustrations and disaffections (Salmela & von Scheve, 2018).

The implications travel further than contemporary populist politics. *Ressentiment* can be seen as a universal feature of human beings because "inferiority feelings are to some degree common to all of us since we all find ourselves in positions which we wish to improve" (Adler, 2013, p. 257). Its intensity, however, is not just an individual affair, but also a function of social structures (Scheler, 1915/1961, pp. 7–8) and dominant ideologies, such as competitive individualism (Sandel, 2020). According to Winnicott (1950, p. 176), in troubled societies, members perceive "the external scene in terms of their own internal struggle, and (they) temporarily allow their internal struggle to be waged in terms of the external political scene." This is how *ressentiment* moves from private to public consciousness and back, particularly in societies where collective problems—social injustice, economic insecurity, corruption of institutions—are perceived, and often framed in individualistic terms (Yankelovich, 1975). Recognizing the mental pain of *ressentiment* and seeking socially-minded approaches to alleviate it, are pressing challenges for public policy officials and scholars of volatile and antagonistic grievance politics.

Acknowledgments

We thank Alexandro Nai, Diego Garzia, Loes Aaldering, Frederico Ferreira da Silva, and Katjana Gattermann for their work on putting together this thematical issue, and for organizing the homonymous workshop in November 2021 which gave us the opportunity to collaborate on this fascinating project. We also thank the workshop participants, the colleagues at the HEPPsinki research seminar at the University of Helsinki, the Oxford Political Psychology Seminar Series, the Hot Politics Lab at the University of Amsterdam, and the anonymous reviewers of *Politics and Governance* who read and commented on drafts of our work. Their constructive and astute suggestions significantly improved our manuscript. A big thanks also to Colin Provost for his generous help in proofreading our work, and the editorial team of *Politics and Governance* for their attention to detail and their diligence. Salmela's work was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 832940).

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Achen, C. H., & Bartels, L. M. (2016). *Democracy for realists: Why elections do not produce responsive government*. Princeton University Press.

Adler, A. (2013). *The science of living*. Routledge.

Aeschbach, S. (2017). *Ressentiment: An anatomy* [Doctoral dissertation, University of Geneva]. Archive Ouverte UNIGE. https://archive-ouverte.unige.ch/unige:103621

Betz, H. G. (1993). The new politics of resentment: Radical right-wing populist parties in Western Europe. *Comparative Politics*, *25*(4), 413–427.

Brader, T. (2006). *Campaigning for hearts and minds: How emotional appeals in political ads work*. University of Chicago Press.

Brader, T., Valentino, N. A., & Suhay, E. (2008). What triggers public opposition to immigration? Anxiety, group cues, and immigration threat. *American Journal of Political Science*, *52*(4), 959–978.

Burla, L., Knierim, B., Barth, J., Liewald, K., Duetz, M., & Abel, T. (2008). From text to codings: Intercoder reliability assessment in qualitative content analysis. *Nursing Research*, *57*(2), 113–117.

Capelos, T. (2013). Understanding anxiety and aversion: The origins and consequences of affectivity in polit-

ical campaigns. In N. Demertzis (Ed.), *Emotions in politics: The affect dimension in political tension* (pp. 39–59). Palgrave Macmillan.

Capelos, T., & Demertzis, N. (2018). Political action and resentful affectivity in critical times. *Humanity & Society*, *42*(4), 410–433.

Capelos, T., & Demertzis, N. (2022). Sour grapes: *Ressentiment* as the affective response of grievance politics. *Innovation: The European Journal of Social Science Research*, *35*(1), 107–129.

Capelos, T., & Katsanidou, A. (2018). Reactionary politics: Explaining the psychological roots of anti-preferences in European integration and immigration debates. *Political Psychology*, *39*(6), 1271–1288.

Capelos, T., Katsanidou, A., & Demertzis, N. (2017). Back to black: Values, ideology and the black box of political radicalization. *Science and Society*, *35*, 35–68.

Capelos, T., Salmela, M., Sullivan, G., & Chrona, S. (2021, May 6–8). *Ressentiment, collective narcissism, and conservation values* [Paper presentation]. Second Helsinki Conference on Emotions, Populism and Polarisation, Helsinki, Finland.

Ciulla, J. B. (2020). Leadership and the power of resentment/*ressentiment*. *Leadership*, *16*(1), 25–38.

Cramer, P. (2015). Understanding defense mechanisms. *Psychodynamic Psychiatry*, *43*(4), 523–552.

Demertzis, N. (2006). Emotions and populism. In S. Clarke, P. Hoggett, & S. Thompson (Eds.), *Emotion, politics and society* (pp. 103–122). Palgrave Macmillan.

Demertzis, N. (2017). Forgiveness and ressentiment in the age of traumas. In W. R. Thompson (Ed.), *Oxford research encyclopedia of politics*. https://doi.org/10.1093/acrefore/9780190228637.013.146

Demertzis, N. (2020). *The political sociology of emotions: Essays on trauma and ressentiment*. Routledge.

Eckstein, H. (1992). *Regarding politics: Essays on political theory, stability, and change*. University of California Press.

Ekman, P. (2004). *Emotions revealed: Understanding faces and feelings*. Henry Holt and Company.

Elster, J. (1999). *Alchemies of the mind: Rationality and the emotions*. Cambridge University Press.

Faludi, S. (2019). *Stiffed: The roots of modern male rage*. William Morrow.

Farkas, J., & Schou, J. (2019). *Post-truth, fake news and democracy: Mapping the politics of falsehood*. Routledge.

Ferro, M. (2010). *Resentment in history*. Polity Press.

Flinders, M., & Hinterleitner, M. (2022). Party politics vs. grievance politics: Competing modes of representative democracy. *Society*. Advance online publication. https://doi.org/10.1007/s12115-022-00686-z

Forgas, J., Crano, W., & Fiedler, K. (Eds.). (2021). *The psychology of populism: The tribal challenge to liberal democracy*. Routledge.

Frijda, N. H. (2004). Emotions and action. In A. S. R. Manstead, N. Frijda, & A. Fischer (Eds.), *Feelings and emotions: The Amsterdam symposium* (pp. 158–173). Cambridge University Press.

Gadarian, S., & Albertson, B. (2012). *Anxious politics: The effect of anxiety of political trust*. Cambridge University Press.

Green, J., Willis, K., Hughes, E., Small, R., Welch, N., Gibbs, L., & Daly, J. (2007). Generating best evidence from qualitative research: The role of data analysis. *Australian and New Zealand Journal of Public Health*, *31*(6), 545–550.

Halsall, R. (2005). Sloterdijk's theory of cynicism, *ressentiment* and "horizontal communication." *International Journal of Media & Cultural Politics*, *1*(2), 163–180.

Hentschel, U., Smith, G. J., Ehlers, W., & Draguns, J. G. (Eds.). (1993). *The concept of defense mechanisms in contemporary psychology*. Springer.

Hochschild, A. R. (2016). *Strangers in their own land: Anger and mourning on the American right*. The New Press.

Hoggett, P. (2018). *Ressentiment* and grievance. *British Journal of Psychotherapy*, *34*(3), 393–407.

Hox, J. J., & Boeije, H. R. (2005). Data collection, primary versus secondary. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 1, pp. 593–599). Elsevier.

Huddleston, A. (2021). Ressentiment. *Ethics*, *131*(4), 670–696.

Huddy, L., Feldman, S., Capelos, T., & Provost, C. (2002). The consequences of terrorism: Disentangling the effects of personal and national threat. *Political Psychology*, *23*(3), 485–509.

Huddy, L., Feldman, S., & Cassese, E. (2008). On the distinct political effects of anxiety and anger. In W. R. Neuman, G. E. Marcus, M. MacKuen, & A. N. Crigler (Eds.), *The affect effect* (pp. 202–230). University of Chicago Press.

Jasper, J. M. (2014). Constructing indignation: Anger dynamics in protest movements. *Emotion Review*, *6*(3), 208–213.

Jost, J. T., Stern, C., Rule, N. O., & Sterling, J. (2017). The politics of fear: Is there an ideological asymmetry in existential motivation? *Social Cognition*, *35*(4), 324–353.

Katsafanas, P. (2022). Group fanaticism and narratives of ressentiment. In L. Townsend, H. B. Schmid, M. Staudigl, & R. Tietjen (Eds.), *The philosophy of fanaticism: Epistemic, affective, and political dimensions* (pp. 157–183). Routledge.

Kazlauskaitė, R., & Salmela, M. (2021). Mediated emotions: Shame and pride in Polish right-wing media coverage of the 2019 European Parliament elections. *Innovation: The European Journal of Social Science Research*, *35*(1), 130–149. https://doi.org/10.1080/13511610.2021.1952551

Kimmel, M. (2017). *Angry white men: American masculinity at the end of an era*. Hachette.

Kinnvall, C., & Svensson, T. (2022). Exploring the populist "mind": Anxiety, fantasy, and everyday populism. *The*

British Journal of Politics and International Relations, *24*(3), 526–542.

Kisić Merino, P., Capelos, T., & Kinnvall, C. (2021). Getting inside "the head" of the far right: Political context and socio-psychological responses. In S. D. Ashe, J. Busher, G. Macklin, & A. Winter (Eds.), *Researching the far right: Theory, method and practice* (pp. 74–91). Routledge.

Kiss, B. (2021). Double ressentiment: The political communication of Kulturkampf in Hungary. *Politics and Governance*, *9*(3), 227–236.

Knauft, B. M. (2018). On the political genealogy of Trump after Foucault. *Genealogy*, *2*(1), Article 4. https://doi.org/10.3390/genealogy2010004

Lazarus, R. S. (1993). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology*, *44*(1), 1–22.

Leiter, B. (2014). *Nietzsche on morality*. Routledge.

León, R., Romero, C., Novara, J., & Quesada, E. (1988). Una escala para medir el resentimiento crítica [A scale to measure critical resentment]. *Revista Latinoamericana de Psicología*, *20*(3), 331–354.

Lerner, J. S., Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science*, *14*(2), 144–150.

Lewis, H. B. (1971). Shame and guilt in neurosis. *Psychoanalytic Review*, *58*(3), 419–438.

Marcus, G. E. (2021). The rise of populism: The politics of justice, anger, and grievance. In J. P. Forgas, W. D. Crano, & K. Fiedler (Eds.), *The psychology of populism: The tribal challenge to liberal democracy* (pp. 81–104). Routledge.

Mason, L., Wronski, J., & Kane, J. (2021). Activating animus: The uniquely social roots of Trump support. *American Political Science Review*, *115*(4), 1508–1516.

Mayer, S. J., & Nguyen, C. G. (2021). Angry reactionary narcissists? Anger activates the link between narcissism and right-populist party support. *Politics and Governance*, *9*(3), 248–259.

Meltzer, B. N., & Musolf, G. R. (2002). Resentment and ressentiment. *Sociological Inquiry*, *72*(2), 240–255.

Michel, E., Garzia, D., Ferreira da Silva, F., & De Angelis, A. (2020). Leader effects and voting for the populist radical right in Western Europe. *Swiss Political Science Review*, *26*(3), 273–295.

Mishra, P. (2017). *Age of anger: A history of the present*. Farrar, Straus, and Giroux.

Moruno, M. D. (2013, September 16–18). *Romantic feelings: Continuities and discontinuities in the history of emotions* [Paper presentation]. Continuities and Discontinuities in the History of Emotions: From the Aftermath of the French Revolution to the Revolutions of 1848, Geneva, Switzerland.

Mudde, C., & Rovira Kaltwasser, C. (2018). Studying populism in comparative perspective. *Comparative Political Studies*, *51*(13),1667–1693.

Nai, A. (2021). Fear and loathing in populist campaigns? Comparing the communication style of populists and non-populists in elections worldwide. *Journal of Political Marketing*, *20*(2), 219–250.

Negra, D., & Leyda, J. (2021). Querying "Karen": The rise of the angry white woman. *European Journal of Cultural Studies*, *24*(1), 350–357.

Nietzsche, F. (1961). *Genealogy of morals*. Cambridge University Press. (Original work published 1885)

Norris, P., & Inglehart, R. (2019). *Cultural backlash: Trump, Brexit, and the rise of authoritarian-populism*. Cambridge University Press.

Nussbaum, M. (2016). *Anger and forgiveness: Resentment, generosity, justice*. Oxford University Press.

Phoenix, D. (2019). *The anger gap: How race shapes emotion in politics*. Cambridge University Press.

Pirc, T. (2018). The collective of ressentiment: Psychopolitics, mediocrity, and morality. In T. Pirc (Ed.), *Participation, culture and democracy: Perspectives on public engagement and social communication* (pp. 111–132). Cambridge Scholars Publishing.

Posłuszna, E. (2019). Inferiority and violence-Islamist terrorism seen against the ressentiment theory. *Toruńskie Studia Międzynarodowe*, *1*(12), 169–179.

Posłuszna, E. (2020). Terrorism of lone wolves. *Safety & Defense*, *6*(1), 109–118.

Posłuszna, E., & Posłuszny, J. (2015). The trouble with ressentiment. *Ruch Filozoficzny*, *71*(4), 87–104.

Reynolds, B. (2004). The nostalgic turn and the politics of ressentiment. *Georgia Educational Researcher*, *2*(1), Article 1. https://doi.org/10.20429/ger.2004.020101

Rico, G., Guinjoan, M., & Anduiza, E. (2017). The emotional underpinnings of populism: How anger and fear affect populist attitudes. *Swiss Political Science Review*, *23*, 444–461.

Roseman, I., & Evdokas, A. (2004). Appraisals cause experienced emotions: Experimental evidence. *Cognition and Emotion*, *18*(1), 1–28.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, *52*(1), 5–18.

Salice, A., & Salmela, M. (2022). What are emotional mechanisms? *Emotions and Society*, *4*(1), 49–68.

Salmela, M., & Capelos, T. (2021). Ressentiment: A complex emotion or an emotional mechanism of psychic defenses? *Politics and Governance*, *9*(3), 191–203.

Salmela, M., & von Scheve, C. (2017). Emotional roots of right-wing political populism. *Social Science Information*, *56*(4), 567–595.

Salmela, M., & von Scheve, C. (2018). Emotional dynamics of right-and left-wing political populism. *Humanity & Society*, *42*(4), 434–454.

Sandel, M. (2020). *The tyranny of merit*. Farrar, Straus, Giroux.

Scheff, T. J. (1994). *Bloody revenge*. Westview.

Scheler, M. (1961). *Ressentiment*. Marquette University Press. (Original work published 1915)

Schreier, M. (2014). Qualitative content analysis. In

U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 170–183). SAGE.
Sharafutdinova, G. (2020). *The red mirror: Putin's leadership and Russia's insecure identity*. Oxford University Press.
Sloterdijk, P. (2010). *Rage and time: A psychopolitical investigation*. Columbia University Press.
Solomon, R. C. (1995). *A passion for justice: Emotions and the origins of the social contract*. Rowman & Littlefield.
Soroko, E. (2014). Identifying indicators of defensive activity in narration about important interpersonal relations. *Current Issues in Personality Psychology*, *2*(3), 149–163.
Sullivan, G. B. (2021). Political reactionism as affective practice: UKIP supporters and non-voters in pre-Brexit England. *Politics and Governance*, *9*(3), 260–273.
Szanto, T. (2018). In hate we trust: The collectivization and habitualization of hatred. *Phenomenology and the Cognitive Sciences*, *19*(3), 453–480.
Szanto, T., & Slaby, J. (2020). Political emotions. In T. Szanto & H. Landweer (Eds.), *The Routledge handbook of phenomenology of emotions* (pp. 478–492). Routledge.
TenHouten, W. (2018). From ressentiment to resentment as a tertiary emotion. *Review of European Studies*, *10*(4), 49–64.
Thomas, S. P. (2003). Men's anger: A phenomenological exploration of its meaning in a middle-class sample of American men. *Psychology of Men & Masculinity*, *4*(2), 163–175.
Thompson, S. (2006). Anger and the struggle for justice. In S. Clarke, P. Hoggett, & S. Thompson (Eds.), *Emotion, politics and society* (pp. 123–144). Palgrave Macmillan.
Turner, J. H. (2007). Self, emotions, and extreme violence: Extending symbolic interactionist theorizing. *Symbolic Interaction*, *30*(4), 501–530.
Vaillant, G. E. (1993). Ego mechanisms of defense and personality psychopathology. *Journal of Abnormal Psychology*, *103*(1), 44–50.
Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice*, *6*(5), 100–110.
Webster, S. W. (2020). *American rage: How anger shapes our politics*. Cambridge University Press.
Wimberly, C. (2018). Trump, propaganda, and the politics of *ressentiment*. *The Journal of Speculative Philosophy*, *32*(1), 179–199.
Winnicott, D. W. (1950). Some thoughts on the meaning of the word democracy. *Human Relations*, *3*(2), 175–186.
Yankelovich, D. (1975). The status of ressentiment in America. *Social Research*, *42*(4), 760–777. https://www.jstor.org/stable/41582863
Žižek, S. (2008). *Violence*. Profile Books.

About the Authors

Tereza Capelos is an associate professor in political psychology and director of the Institute for Conflict Cooperation and Security at the Department of Political Science and International Studies at the University of Birmingham. She is past president of the International Society of Political Psychology, co-convenor for the Political Psychology Standing Group of the European Consortium for Political Research, and co-editor of the Palgrave Series in Political Psychology. Her research examines reactionary orientations and *ressentiment*ful affect as psychological expressions of grievance politics, and the role of uncertainty and affective ambivalence in establishing trust during national and international crises and tensions.

Mikko Salmela is an adjunct professor of practical philosophy and a member of the Helsinki Hub on Emotions, Populism, and Polarisation (HEPPsinki) at the University of Helsinki, and an associate professor at the Centre for Subjectivity Research at the University of Copenhagen. His main research interests are in empirically informed philosophy of emotion, philosophical and political psychology, and philosophy of sociality.

Gabija Krisciunaite holds a first-class honors for the joint Master of Arts degree in economics and international relations at the University of Aberdeen, and a first-class honors Master of Science degree in political psychology of international relations at the University of Birmingham. She researches the collective character of *ressentiment* and its link to narcissism and reactionism.

Politics and Governance (ISSN: 2183–2463)
2022, Volume 10, Issue 4, Pages 396–410
https://doi.org/10.17645/pag.v10i4.5790

Article

Online Trolls: Unaffectionate Psychopaths or Just Lonely Outcasts and Angry Partisans?

Monika Verbalyte [1,*], Christoph Keitel [1], and Krista Howard [2]

[1] Interdisciplinary Center for European Studies, Germany
[2] Department of Psychology, Texas State University, USA

* Corresponding author (monika.verbalyte@uni-flensburg.de)

Submitted: 13 May 2022 | Accepted: 10 September 2022 | Published: 30 December 2022

Abstract
The main objective of the article is to attempt to provide a more sociological explanation of why some people attack and insult others online, i.e., considering not only their personality structure but also social and situational factors. The main theoretical dichotomy we built on is between powerful high-status and low-on-empathy "bullies" trolling others for their own entertainment, and people who are socially isolated, disempowered, or politically involved, therefore feel attacked by others' beliefs and opinions expressed online, and troll defensively or reactively instead of primarily maliciously. With an MTurk sample of over 1,000 adult respondents from the US, we tested these assumptions. We could confirm that there are two categories and motivations for trolling: for fun and more defensive/reactive. Further, we checked how strongly precarious working conditions, low social status, social isolation, and political as well as religious affiliation of the person increase or decrease the probability of trolling as well as enjoyment levels from this activity. We controlled for personality traits, social media use and patterns, as well as sociodemographic factors. We could confirm that political identities and religiosity increase the likelihood of, but not the enjoyment of trolling; however, socio-economic factors do not have the same differentiating effect.

Keywords
negative politics; online deviance; political affiliation; powerlessness; social media; trolling; USA

Issue
This article is part of the issue "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" edited by Alessandro Nai (University of Amsterdam), Diego Garzia (University of Lausanne), Loes Aaldering (Free University Amsterdam), Frederico Ferreira da Silva (University of Lausanne), and Katjana Gattermann (University of Amsterdam).

© 2022 by the author(s); licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Social media is often seen as a polarizing, if not toxic, environment. In part, this heated online climate might be attributed to constant conflicts with so-called internet "trolls." It is still not fully clear why some people find it appropriate and even fun to attack and insult others online. The most widespread explanation, offered mainly by psychological studies, is that those who engage in such behavior are aggressive individuals with low impulse control, psychotic personality traits (Buckels et al., 2014; Lopes & Yu, 2017), and a lack of empathy (Sest & March, 2017). However, this research rarely considers how the social environment and living conditions of trolls—such as stress level, precarious working conditions, social isolation, economic deprivation, and other hardships—may influence their engagement in different kinds of social and online deviance. Also, scholars have not yet examined how trolling behavior is influenced by political identities, i.e., those known to strongly increase affective polarization by driving people to react aggressively to their opponents, especially within the anonymous internet space. On this subject, one qualitative interview study recently found that the most aggressive

people online are not only extreme in their political attitudes but also disadvantaged in their social lives, such that they seek needed attention, empowerment, and confidence from hostile online activities (Bail, 2021).

The goal of this study is to theoretically contrast and comprehensively analyze these different modes of trolling, focusing on the personality traits of individuals who engage in such behavior and the socio-political factors that may impact their decision to do so. We see this phenomenon in the broader context of current developments: rising affective political and social polarization and negative partisanship in both the US and Europe (Reiljan, 2020), the increasing negativity of electoral campaigns (Nyhuis et al., 2020), and the general incivility of political discourse (Gidron et al., 2019). Very often, this augmented hostility is related to populist and extremist (especially far-right) parties, whose candidates are more prone to use attack politics and harsh campaigns (Nai, 2021), and more often loath and are loathed by the partisans of other parties (Harteveld et al., 2021).

Against this background, a puzzle arises. While dark personality traits allegedly underlie trolling behavior, the phenomenon also seems to represent a reactive response to felt powerlessness. Former US President Donald Trump can serve as an illustrative example here. Trump can be described as a real "bully": narcissistic, aggressive, power-obsessed (Nai & Maier, 2018), and seemingly liked by individuals with similar personality profiles (Nai et al., 2021). It is logical to expect that people like Trump and his supporters are also active and spread hostility online. Yet so-called "rednecks" and other populist supporters are also often seen as social "losers": individuals whose economic situation objectively worsened in the last decades (Gidron & Hall, 2017), who feel misunderstood, unjustly treated, disempowered, and excluded due to their beliefs (Abts & Baute, 2021; Hochschild, 2016), and who fuel *ressentiment* and reactionarism (Salmela & Capelos, 2021). This negative emotionality is channeled not only through populist support, but also vented on social media, for example by trolling out-party supporters (Bail, 2021)—thereby further fueling affective polarization (Bulut & Yörük, 2017).

Considering these factors, this article seeks to take a more comprehensive look at trolling behavior. First, we clarify the definition of trolling and differentiate between trolling for fun and more defensive/reactive trolling. Second, we review the possible explanations of online trolling behavior, including the more sociological and political ones. We continue by performing explanatory and confirmatory factor analyses (EFA and CFA, respectively) to test the assumed differences between concepts of and motivating factors behind trolling, as well as structural equation path modeling to show their relationship. We then conduct regression analyses in order to introduce a wider range of explanatory and control variables into the equation. Finally, we conclude by presenting our results and reflecting on the limitations of our study.

2. Definition of Trolling

The term "trolling" has become a catchphrase to describe a range of online deviant behaviors. It is crucial, therefore, that we differentiate between trolling and other forms of anti-social, or dark, online activities, the most similar of which are hate speech, online harassment, and cyberbullying (Hardaker, 2010; Shachaf & Hara, 2010). All four must be clearly distinguished from impoliteness, online incivility, offensive, derogatory, and abusive language, or so-called "flaming." The latter depict forms of discursive hostility (Andersen, 2021), or hostile ways of presenting content—not the behavior itself. Accordingly, uncivil language could be used in connection with all forms of online hostility, albeit for different purposes and with different intentions. What differentiates trolling from other forms of hostile online behavior is, first and foremost, the target. There is a pre-existing relationship between the cyberbully and the victim(s), so damage is caused to the specific target (Fichman & Sanfilippo, 2016). Harassment is related to real-world stalking and is often sexual in nature (May et al., 2012), and hate speech is a form of verbal violence directed at particular groups, especially on the basis of race, religion, or sexual orientation (Paasch-Colberg et al., 2021). By contrast, in the case of trolling the target does not need to be predefined, and the network of trolling is usually wide, not limited as it is with cyberbullying (Hardaker, 2010).

All in all, while trolling has become a widely known internet phenomenon, its definition remains blurry due to the variety of practices it encompasses. Trolling activities range from mocking others for self-amusement to disrupting online communities (Pao, 2015). In general, trolling can be defined as the act of posting/sharing content or comments that may irritate or cause conflict among those who receive it, and/or starting a circular discussion that deliberately provokes an aggressive response by the targets of the trolling (Hardaker, 2010). Fichman and Sanfilippo (2015) regard online trolling as a deviant and antisocial behavior in which the user acts provocatively against normative expectations. However, when we look at this phenomenon with the aim to understand the desired outcomes and intentions of trolling, the picture becomes rather vague. From this perspective, we encounter enjoyment and thrill-seeking as well as revenge and self-gratification (Cook et al., 2018; Fichman & Sanfilippo, 2015; Quandt et al., 2022; Shachaf & Hara, 2010). This article seeks to manage the complexity of the trolling dynamic by categorizing trolling into two main types: (a) trolling for fun and (b) defensive/reactive trolling.

Trolling as a form of entertainment has been widely documented (Buckels et al., 2014; Shachaf & Hara, 2010). Hardaker (2010) explains that trolls aim to cause disruption and trigger unpleasant feelings in their victims for the sake of their own amusement. The trolling behavior is therefore associated with positive emotions, such

as the pleasure that comes from attracting attention. Yet the internal trigger of boredom also appears to play an important role in this context (Shachaf & Hara, 2010).

In contrast, defensive/reactive trolling is related to negative emotions and represents a reaction to another person's actions. In this vein, Cheng et al. (2017) found that exposure to previous trolling behavior by others can increase the users' likelihood of engaging in trolling behavior themselves. Furthermore, witnessing how another group member is subjected to trolling may motivate a third party to step in and troll back (Hopkinson, 2013). Another possibility is the so-called "white and gray hat trolls"—or people who not only troll back, but also "troll the trolls" for the sake of an alleged higher cause, i.e., to direct public attention to certain issues (Matthews & Goerzen, 2019). In this sense, trolling might be motivated by genuine beliefs and in-group identities, and not only by feelings of boredom or the expectation of fun. People who engage in this kind of trolling might feel threatened, attacked (Liu et al., 2018), and vengeful (Shachaf & Hara, 2010). They may respond spontaneously and impulsively, without the intent to harm (Hardaker, 2010), and later regret their actions. To sum up, this trolling is defensive/reactive in many respects: It is a neurotic or moody reaction to the provocation of another, as well as a more considerate response strategy based on strong beliefs and calculations.

3. Explanations of Trolling Behavior

The different types of trolling already provide some hints as to the underlying reasons behind such behavior. This section reviews the known explanatory factors for trolling, focusing on those related to defensive/reactive trolling. To gain access to a broader pool of potential explanations, we consulted a wide range of studies on online aggressiveness, deviance, and criminality in general. We start by reviewing what Cook et al. (2018) classified as the internal triggers (personality traits) and circumstantial factors (mainly, the social media environment) of trolling behavior. While these are the most commonly considered factors with regard to trolling, they are rather narrow in their explanatory power. After that, we present what we believe to be the most important social and political determinants of defensive/reactive trolling.

3.1. Personality Structure

Among scholars, personality structure appears to be the most popular and attractive explanation of trolling behavior. The dark personality types—Machiavellianism, psychopathy, sadism, and narcissism—are distinguished by remorseless, impulsive, thrill-seeking, and socially offensive behavior (Buckels et al., 2019; Jones & Paulhus, 2011), including trolling. Psychopathy is related to low self-control and low empathy, as well as a lack of respect for social norms and conventions (Foster & Trimm, 2008). The absence of inhibitory mechanisms permits people with this trait to engage in a wide range of antisocial activities (Craker & March, 2016; Hare, 2006; Jones et al., 2011). Sadistic psychopaths also believe that it is totally acceptable to manipulate and use others, for example by blackmailing and bullying popular individuals (Buckels et al., 2014; Lopes & Yu, 2017). Like psychopaths, narcissists and Machavellianists have a distorted view of their own self-importance and do not hesitate to use manipulation and deception for self-enhancement and to protect their self-esteem, although their behavior is less motivated by pleasure-seeking or harming others and is more self-serving than that of psychopaths (Campbell & Miller, 2011; Jakobwitz & Egan, 2006). Our data set, however, only enables us to measure empathy, which has already been shown to negatively relate to both psychopathy and trolling (Sest & March, 2017).

Our analysis does include the Big5 inventory, of which low levels of conscientiousness and agreeableness are associated with psychopathy and a higher likelihood of enjoying trolling. High levels of extraversion, neuroticism, or openness also have a positive impact on trolling (Gylfason et al., 2021); however, extraversion and openness might well indicate an outgoing and self-confident personality and an interest in communicating in various different ways, rather than malignant intentions. Moreover, neuroticism, low levels of agreeableness, and low levels of conscientiousness are correlated with a specific kind of social media use—that is, the kind that seeks to gain from social media the attention and social support people might lack in their offline lives (Seidman, 2013), so also possibly defensive trolling. Another known personality-related explanation of online deviance has to do with the difficulties some people experience in regulating their own emotions, and the ramifications of their anger (Toro et al., 2020). This negative affect is known to spiral, completely consume the person, and reduce the person's inhibitions, prompting an urge to engage in corrective action towards the culprit (the person perceived to have caused the anger; Agnew, 1995). In this sense, it could also be related to the troll-back reaction.

Thus, we expect empathy to have a strong impact in that it can decrease a person's enjoyment of trolling for fun (but not defensive trolling). However, the contradictory reasoning and differing results of the existing studies on the Big5 hinder us from formulating any clear expectations on this relation. Because our survey provides no measure for emotional regulation as a personality trait, we count the motivations to troll related to anger and other negative emotions as situational, and therefore as reactive trolling.

3.2. Social Media Environment

With the recent development of new technologies, trolling opportunities have become more available. Many online platforms, especially those supported by social media, provide relatively open spaces for a large

number of individuals to engage in genuine, sincere interpersonal communication and debate. Trolls actively prey on such individuals by posting controversial and inappropriate messages to derail discussions into pointless tangents (Herring et al., 2002). Little to no moderation and the option to remain anonymous means that violations of conversational norms often go unpunished, or at least give that impression—which, in turn, invites deception, controlled self-presentation, and decreased self-control among some users (Hardaker, 2010). The ability to remain incognito or to disguise their identities gives trolling individuals an advantage over their victims, who often openly share personal information about themselves, thereby inadvertently inviting ad hominem attacks. Another known effect of anonymity is its deindividuating effect on users: Decreasing the salience of individual identity and increasing that of social/group identity deflects personal responsibility and enables deviant behavior (see social identity model of deindividuation effects by Spears, 2017, and our notes on group-driven trolling in the last passage of Section 2).

We also follow the logic of situational action theory, which addresses both the personal and environmental factors that tend to predominate in the offending individual (Li et al., 2022). Since the impact of the online environment is constant for everyone, what creates the difference is how individuals approach and use that environment. In this regard, time appears to be one key factor: Scholarship tends to show a strong association between time spent online and deviant behavior online (Lee, 2018). What about different kinds of trolling—Do they seem to align with different usage patterns regarding the online environment? On the one hand, more people show dark personality traits online than offline (Nevin, 2015) and hostile political discussions are mainly caused by status-driven individuals (Bor & Petersen, 2021). On the other hand, online attacks are frequently driven by people who feel their lack of status and want to boost it (Bail, 2021). In this respect, it seems that online environments do not feed into any specific type or types of trolling; they appear to promote trolling of all kinds. Also, we know that interactions between ideologically opposed users are significantly more negative than like-minded ones (Marchal, 2021), which means that the composition of online networks matters. Once again, this seems to present more opportunities for trolling in general—of all kinds.

How people shape social comparison on social media, however, is differently related to different sorts of trolling. Research suggests that mostly psychopathic people engage in upward social comparison, or comparison with people who are better off (Lopes & Yu, 2017); others who want to enhance their social status make a downward comparison, since upward comparison for them would only increase their anxiety, low self-esteem, insecurity, and loneliness (Bonnette et al., 2019; Howard et al., 2019). We would expect defensive trolling to be more related to downward online comparison, and enjoyment trolling to be more related to upward online comparison.

3.3. Social Status and Identity

Independently of a user's propensity to anger ramifications, experiencing online hostility from others increases the likelihood that one will respond in a trolling manner (Liu et al., 2018). This could be triggered not only by vengefulness or a lost temper but also by the feeling that social norms allow for this kind of behavior (Sentse et al., 2007)—a feeling encouraged by the relative freedom and anonymity of online space, as we noted in Subsection 3.2. The inhibition of social controls is even stronger in group settings: Research suggests that exposure to trolling does not have to be experienced directly by individuals but may gear itself more towards group identities (see also social identity model of deindividuation effects theory in Subsection 3.2. Trolling out-group and in-group members who violated some inner rules also helps to establish community boundaries (Rafalow, 2015).

In that sense, partisanship represents a form of defensive/reactive trolling, even if it sometimes goes beyond mere revenge and represents some form of ideological impetus (Fichman & Sanfilippo, 2016). Driven by certain political beliefs or ideologies, trolls seek to draw attention to the social issues they care about (Sanfilippo et al., 2017), or they use social media as communication and mobilization channels through which to find allies and/or to legitimate and inspire a group action (Flores-Saviaga et al., 2018). Political trolling practices can also include baiting ideological opponents into arguments through coordinated behavior and conscious attack, in order to spam adversary online platforms and individuals (Frischlich et al., 2021; Sanfilippo et al., 2017). However, in this case, the action is done in response to the troll's perception of a group threat, rather than for personal fun. A somewhat different function could be ascribed to another identification: religion. Religious convictions are often politicized and important in defensive/reactive trolling; yet they are also sometimes presented as moral convictions that deter the holder of those convictions from deviance (Adamczyk, 2012). In that case, the person might feel shame or guilt when trolling, rather than enjoyment.

General socio-economic factors also belong to the important determinants of social deviance. Most of the studies relate to poverty, low levels of education, and destructive family relations to anti-social behavior (Cioban et al., 2021; Hagan & Foster, 2001). Uncivil online behavior has been even more strongly linked to a generally low or reduced social status (for example, due to a recent divorce or unemployment) and subjective feelings of social isolation (such as lack of social support) and economic deprivation. The loneliness of people with these characteristics leads them to attach greater importance to other social identities, such as political

partisanship (cf. Salmela & von Scheve, 2018) and to vent online more often than other individuals (Bail, 2021), which in turn ostensibly enhances their power and social status (Juvonen & Graham, 2014). The feeling of injustice might be more relevant than the objective socio-economic situation; thus, a person could have a fairly solid social standing and still lack the experience of social mobility (cf. Hochschild, 2016; see also Salmela & von Scheve, 2018) or feel deprived of what they legitimately deserve (Agnew, 1995; Cioban et al., 2021). This frustration causes anger and aggression as well as anxiety, depression, and general irritability (Aseltine et al., 2000), which leads us to believe that it also belongs to a group of factors causing defensive trolling.

To summarize, we expect to find more defensive/reactive trolling among people with strong political identities than among other individuals. However, we cannot formulate a clear hypothesis for religious identity, which might also inhibit trolling. Second, we suggest that people with low socio-economic status, precarious working conditions, and relatively low levels of social support are more prone to troll than others. On the other hand, an elevated social status appears more related to trolling for fun, since having a successful career and a position of power might also signal the presence of a dark personality (i.e., given the deep disregard for others and strong will to dominate that tend to characterize individuals with dark personality traits; Festl & Quandt, 2013; see also Section 3.1).

4. Research Data and Design

To test our theoretical model, we used survey data of 1,314 individuals gathered by Texas State University on MTurk (Amazon Mechanical Turk). The sample consists of 51.26% men and 48.29% women (the remaining respondents reported as transgender or gender non-conforming), 68.57% of whom self-identified as white, 40.79% as Democrats, and 26.64% as Republicans. The mean age of respondents was 35.74 years. In comparison to US census data, the sample overrepresents young people and Democrats; however, this is fairly common in online samples. While there are also some deviations for race, this might be based on the different race categories used and the fact that not all MTurk workers come from the US. Although scholars often express reservations about the quality of MTurk surveys, they still seem to be preferred over purely student samples (Pew Research Center, 2016).

Since the original purpose of the survey differs from the aim of our study, it does not cover all the variables that follow from our theoretical considerations, yet still provides a wide range of variables that could help to answer some of our research questions. In this sense, our study could be seen as a secondary data analysis. For this reason, we also present our study as rather explorative, even though it is generally led by theoretical considerations and expectations.

In analyzing these variables, we first present the available items related to trolling and the motivation to troll. We then draw on both EFA and CFA to group these into meaningful variables. Later, we present the expected links between those variables and test them with structural equation path modeling. To include more variables than would be feasible with the path model, we continue with regression analyses. Analyses were conducted with Stata15 and R package lavaan.

4.1. Trolling and Trolling Motivations

For our dependent variables, we decided to go with the items that would represent the greatest breadth of possible trolling and attacking behaviors and hypothesized that they could be clustered into two categories. In the first category, we grouped the items concerned with the extent to which one enjoys debating to upset or irritate others and to troll. These items were expected to represent the enjoyment of trolling. In the second category, we grouped together items related to the respondent's likelihood to make comments, create posts, send people "shock websites" to upset others, and whether the respondent enjoys posting with the intent to annoy others. We used the items in this second category to measure the respondent's likelihood to troll, since they do not specifically mention enjoyment and therefore probably also represent defensive/reactive trolling (for a more detailed description of variables, see the Supplementary Material).

This structure of variables could already be confirmed by means of a correlation matrix (Figure A1 in the Supplementary Material). Since all of the items represent anti-social and undesirable behaviors, they did not show a normal distribution and had to be dichotomized for further analyses (similar to Howard et al., 2019), i.e., all respondents who did say more than "extremely unlikely" or "strongly disagree" (for descriptives, see Figure A2 in the Supplementary Material). While this is a very strict exclusion criterion for the circle of trolls, it was done in order to capture measurements for accidental trolling behavior. EFA renders the model with two factors (for Table 1, Promax rotation was used because the respondent's likelihood to troll also partially includes trolling for fun, so these two sorts of trolling cannot be treated as independent from one another). CFA confirms this structure only with acceptable fit (Table A1 with loadings in the Supplementary Material), but we decided not to further modify our model due to the strong theoretical assumptions and other empirical evidence for this model. We created two continuous variables from the corresponding EFA factor scores, which ranged from 0 to 1. The pressing question, however, is how keen people are to admit that they are trolling, and how reasonable our partition of trolling in these two groups is. We present some validity checks to support our claims and data (see validity checks in the Supplementary Material), which confirm that both our numbers of trolls

Table 1. Factor loadings of trolling items.

	Likely to troll	Enjoy trolling
Comment to upset	**0.8007**	0.1280
Offensive posts	**0.9411**	0.0059
Send shocking websites	**0.8669**	0.0240
Post to annoy	**0.4228**	0.2661
Enjoy trolling	0.2742	**0.5294**
Enjoy debating to upset	−0.0064	**0.8909**

Notes: With Promax rotation, normalized.

as well as our inquiry into different trolling motivations are justified.

To merge deeper into these possible motivations, we look at the survey questions that address why respondents troll or how they felt while doing it. The questions address different emotional, mental, and social states of the people: "When you are commenting or posting in order to upset or irritate others, to what extent are you… Highly Stressed, Tired/Fatigued, Bored, Depressed, Anxious, Feeling Attacked, Lonely, Silly, Annoyed" and "When commenting on others' posts with the intent to upset or irritate others, to what extent do you feel… Powerful, Courageous, Intelligent, Levelheaded, Happy, Embarrassed, Devious, Superior, Cruel, Protected, Guilty, Confident, Fearful, Defensive." All of the answers are measured on a scale of 0 to 100. Since named motivations had a strong relation to trolling behavior in terms of distribution, we also dichotomized these items. Thus, everything higher than 10 was considered to include this motivation.

We used these items to build patterns of trolling which Maltby et al. (2016) called implicit trolling theories. First, we proceeded with EFA. The exclusion of the items Devious, Defensive, and Fearful, has rendered a good five-factor solution (Table 2). Already in the correlation matrix (Figure A3 in the Supplementary Material), excluded variables showed low correlations with other items or did not constantly load on any specific factor. The CFA also provided a good fit for this five-factor solution (Table A3 in the Supplementary Material), with the factors being Powerful and Happy, Silly and Bored, Attacked and Annoyed, Stressed and Anxious, and Embarrassed and Guilty. Here we also created continuous variables from corresponding EFA factor scores, which ranged from 0 to 1.

The literature we reviewed mentioned all of the motivation groups as possible motives for trolling. Here, we assumed that Powerful and Happy and Silly and Bored would be more related to trolling for fun, whereas Attacked and Annoyed (due to the defensive,

Table 2. Factor loadings of trolling motivations.

	Stressed and Anxious	Attacked and Annoyed	Powerful and Happy	Embarrassed and Guilty	Silly and Bored
Highly Stressed	**0.6849**	0.2568	−0.0167	0.0897	−0.1395
Tired/Fatigued	**0.6835**	0.0910	0.0831	−0.0008	0.0395
Depressed	**0.7731**	−0.0309	0.0385	0.0266	0.1032
Lonely	**0.6415**	−0.1190	0.1143	−0.0192	0.2980
Anxious	**0.8053**	0.1107	−0.0038	0.0257	−0.0383
Feeling Attacked	0.1606	**0.6424**	0.0424	0.1236	−0.0248
Annoyed	0.0906	**0.7981**	−0.0070	−0.0192	0.1166
Happy	−0.0298	−0.0458	**0.9403**	−0.0014	0.0078
Powerful	0.0475	−0.0155	**0.8744**	0.0383	−0.0000
Superior	0.0290	0.0447	**0.8112**	0.0396	0.0557
Confident	−0.0612	0.1128	**0.9076**	−0.0813	0.0164
Courageous	0.1219	−0.0596	**0.8657**	0.0425	−0.0628
Intelligent	−0.0322	0.0539	**0.9329**	−0.0452	−0.0015
Levelheaded	−0.0200	0.0331	**0.9054**	−0.0301	−0.0358
Protected	0.0823	−0.0995	**0.7474**	0.0794	0.0994
Embarrassed	0.1253	−0.0580	0.1400	**0.7165**	−0.0854
Guilty	0.0399	0.0048	−0.0914	**0.8997**	0.0180
Cruel	−0.0989	0.0830	−0.0105	**0.8575**	0.0739
Bored	0.0761	0.1061	0.0899	−0.0106	**0.7034**
Silly	0.2033	0.0039	0.1630	0.0416	**0.4881**

Notes: With Promax rotation, normalized.

possibly spontaneous reaction), Stressed and Anxious (in order to get attention and overcome one's insecurities), Embarrassed and Guilty (due to the inhibiting impact of social norms after the spontaneous outburst) would be indicators of defensive/reactive trolling. Our hypothesized relations do not mean that there could be no cross-loadings. For example, it is possible to feel some *Schadenfreude* (pleasure at another person's misfortune) even while trolling defensively; but since this is not the primary motivation for this behavior, the relation should be much less incisive. These cross-loadings we first and foremost expect between trolling for fun motivations and likelihood to troll, since defensive/reactive trolling could not be clearly extracted from variables on trolling behavior. On the other hand, enjoyment of trolling should clearly only be related to fun motivations (for the main hypothesized model, see Figure 1).

4.2. Other Independent Variables

For further analyses, we will conduct multiple linear regressions since further buildup of the path model would make it too confusing. Also, it would be hard to argue for some paths as mediations or causal links since we only have cross-sectional data.

We start with social status variables (more detailed description in the Supplementary material). Occupation includes jobs of different skill and status levels. Precarious Working Situation differentiates between people having one stable job, and those forced to juggle a few part-time positions. The Interpersonal Support Index is a proxy of available support networks, the opposite of social isolation. Another group of main explanatory variables is identity. For Political and Religious Affiliation, we included both the strength and ideological direction of affiliation. To control for political involvement, we took people's self-assessed online political engagement or Percentage of Political Posts.

The rest of the variables represent controls. For aspects related to personality, we used the Big5 inventory and Empathy. For social media environment, we used the measure of Heavy Social Media Use and two additional variables of how people use them and the composition of their social networks. The former is measured through Upward and Downward Online Comparison, the latter with the Percentage of (Dis)Similar Groups in which they are involved. Finally, we also added a few important socio-demographic variables: Gender, Age, Race, and LGBTQAI+.

5. Results

First, we test the main measurement and path model (see Figure 2). It shows a good fit between our hypothesized model and our data (CFI = 0.951; TLI = 0.9445; RMSEA = 0.064). The measurement part confirms our previously presented results: Factor loadings are high for all of the relations discovered through EFA. The new part is the hierarchical model of motivations where five latent groups are further located under two groups of motivations, and high loadings confirm that the categories Embarrassed and Guilty, Attacked and Annoyed,

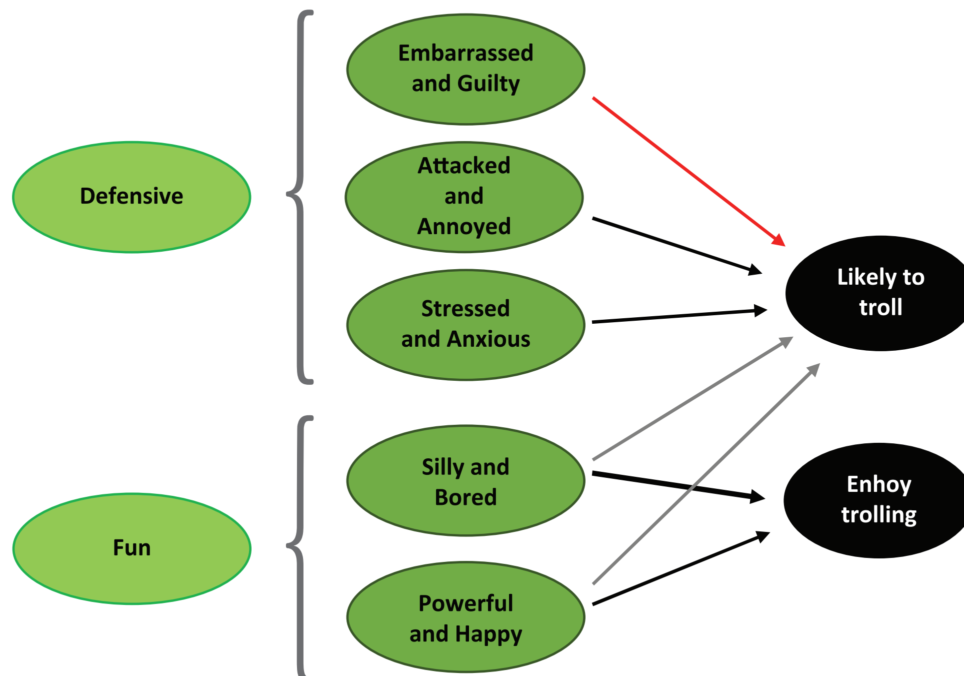

Figure 1. Tested measurement and path model. Notes: All the presented variables are latent variables; part of the measurement model with manifest variables is excluded in order to not overload the figure; different figure colors represent different variable groups; black arrows represent hypothesized positive relations, red ones hypothesized negative relations, and gray possible positive relations.

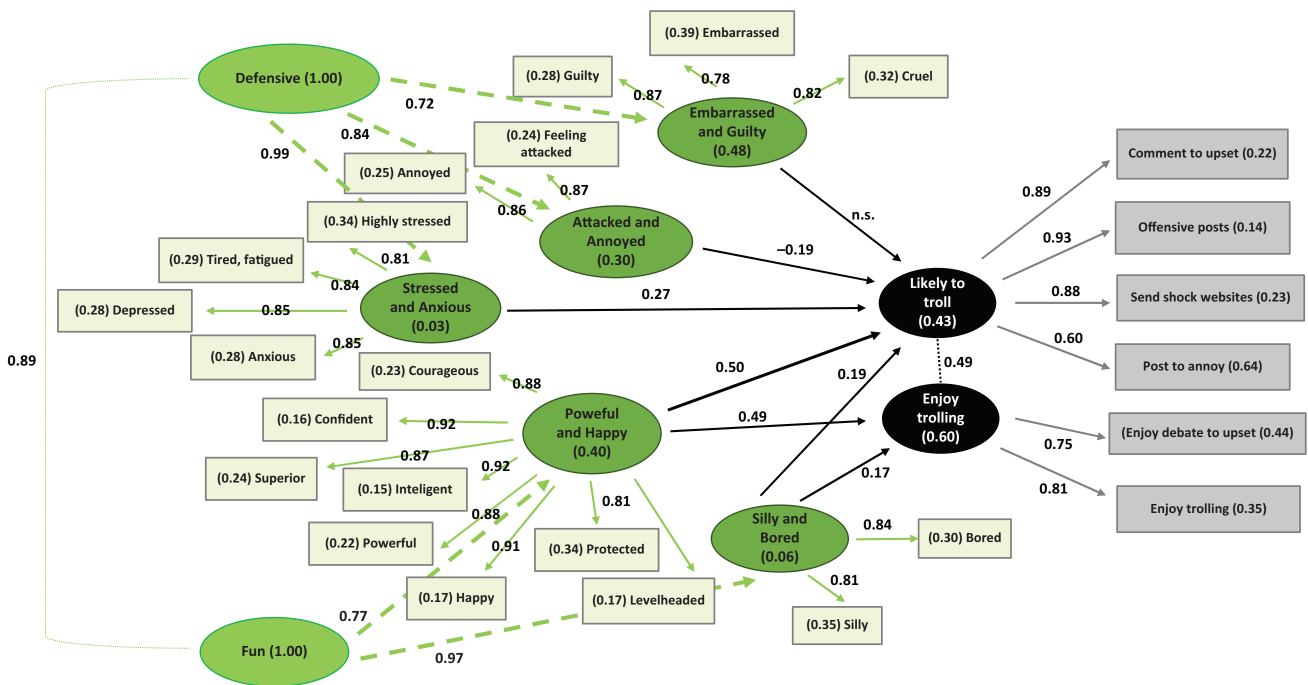

Figure 2. Main measurement and path model. Notes: Near variable names variances are depicted in parentheses; black and gray colors represent trolling; green colors represent motivations to troll; gray arrows are loadings of trolling behaviors on latent trolling variables; green arrows are loadings of specific motivations on latent motivation variables; green dashed lines are loadings of latent motivation variables on two main groups of motivations; black straight lines are regression coefficients; dotted lines (green for motivations, black for trolling behavior) are correlations.

and Stressed and Anxious belong to defensive motivations, whereas Silly and Bored and Powerful and Happy correlate to trolling for fun.

As for our path model, we see that Powerful and Happy as well as Silly and Bored have a significant positive relation with one's enjoyment of trolling and, as expected, one's likelihood to troll. Defensive trolling motivations demonstrate different relations to trolling. Thus, Stressed and Anxious has a positive influence on one's likelihood to troll, whereas—contrary to expectations—Embarrassed and Guilty has no significant relation to trolling behavior and Attacked and Annoyed has a negative coefficient for trolling. As for Embarrassed and Guilty, a nonsignificant relation was as plausible as a negative one. Indeed, in a sense, this feeling does not decrease or increase trolling: It is only felt during or after the act of trolling. As for Attacked and Annoyed, a negative coefficient sign emerges only after controlling for Powerful and Happy and Stressed and Anxious, whereas the binary correlation with trolling is positive (see Figure 5A in Supplementary Material). Thus, it seems that if the person is stressed, or feels self-confident and superior, attacking them does not serve as an additional driver for trolling; rather, it decreases the probability of a trolling incident. Possibly, this is due to a negative psychological effect—i.e., the attack deepens the person's anxiety or partially diminishes their courage to troll.

To control for a wider range of variables and possible cross-loadings between motivations and trolling, we now turn to regressions (see Table 3). Models 1a and 1b just confirm path model results. Models 2a and 2b demonstrate the influence of the main explanatory variables. From our social status variables, interpersonal support demonstrates the strongest impact. Yet we do find strong negative associations between this factor and enjoyment of trolling and likelihood to troll, although we have hypothesized that its impact is limited to general trolling. The explanation for this might be that trolling is not only more likely, but also more enjoyable if a person is bored and alone due to the excitement delivered by the act itself, and the satisfaction of not having to deal with one's lack of meaningful social connections. We also see a strong and positive correlation between the enjoyment of trolling and some occupations—not necessarily higher-status ones. This might simply indicate that, in comparison to unemployment, any job will increase one's feelings of power and therefore one's enjoyment of trolling. The likelihood to troll does not significantly correlate to any single occupation, but having a precarious position is a significant predictor of trolling behavior, even though it reduces the probability of trolling. When we also compare this result to the negative binary link between unemployment and trolling (Figure A4 in the Supplementary Material), we may conclude that having the lowest levels of social status (precariously employed or unemployed) is so disempowering that the person does not even try to gain a feeling of empowerment through trolling.

Table 3. Regression analyses on one's enjoyment of trolling and likelihood of troll.

	(1a) Enjoy trolling	(2a) Enjoy trolling	(3c) Enjoy trolling	(1b) Likely to troll	(2b) Likely to troll	(3b) Likely to troll
Feeling while trolling: Powerful and Happy	0.489*** (0.039)	0.392*** (0.042)	0.331*** (0.042)	0.515*** (0.037)	0.370*** (0.040)	0.302*** (0.037)
Feeling while trolling: Silly and Bored	0.110* (0.065)	0.060 (0.067)	0.017 (0.066)	0.161*** (0.057)	0.092* (0.059)	0.052 (0.057)
Feeling while trolling: Attacked and Annoyed	0.025 (0.055)	0.071 (0.056)	0.100* (0.055)	−0.145*** (0.043)	−0.089** (0.042)	−0.042 (0.041)
Feeling while trolling: Stressed and Anxious	−0.009 (0.068)	−0.091 (0.071)	−0.110 (0.068)	0.205*** (0.057)	0.116* (0.058)	0.062 (0.054)
Feeling while trolling: Embarrassed and Guilty	−0.011 (0.049)	−0.034 (0.048)	−0.015 (0.047)	0.009 (0.041)	−0.021 (0.041)	−0.021 (0.039)
Precarious working position (ref. no such position)		−0.020 (0.063)	0.007 (0.064)		−0.036* (0.039)	−0.012 (0.038)
Occupation (ref. unemployed)						
Manager		0.101* (0.046)	0.060 (0.045)		0.044 (0.036)	0.020 (0.037)
Professional		0.088 (0.040)	0.061 (0.038)		0.029 (0.032)	0.021 (0.032)
Clerk/asisstant professional		0.112** (0.044)	0.088* (0.042)		0.017 (0.035)	0.006 (0.034)
Skilled worker		0.050 (0.057)	0.026 (0.056)		0.055 (0.048)	0.033 (0.044)
Unskilled worker		0.067* (0.074)	0.061 (0.070)		0.011 (0.050)	−0.000 (0.053)
Other occupation		0.027 (0.080)	0.017 (0.072)		0.008 (0.061)	0.005 (0.054)
Interpersonal support		−0.176*** (0.002)	−0.115** (0.002)		−0.140*** (0.001)	−0.038 (0.002)
Political affiliation strength		−0.017 (0.014)	−0.019 (0.013)		0.001 (0.010)	0.001 (0.010)
Political affiliation (ref. no pol. affiliation)						
Democrat		0.008 (0.053)	0.041 (0.051)		0.083 (0.038)	0.136** (0.038)
Republican		0.045 (0.055)	0.059 (0.053)		0.104* (0.041)	0.123** (0.040)
Other political affiliation		0.015 (0.052)	0.042 (0.050)		0.054 (0.037)	0.088* (0.037)
Religiosity (strength)		−0.012 (0.010)	0.020 (0.010)		0.027 (0.008)	0.069* (0.008)
Religion (ref. nonreligious)						
Christian		0.048 (0.041)	0.053 (0.040)		−0.032 (0.032)	−0.032 (0.030)
Hindu		0.085* (0.047)	0.107** (0.054)		0.062 (0.044)	0.050 (0.044)
Spiritual		0.027 (0.060)	0.039 (0.062)		−0.018 (0.040)	−0.016 (0.042)

Table 3. (Cont.) Regression analyses on one's enjoyment of trolling and likelihood of troll.

	(1a) Enjoy trolling	(2a) Enjoy trolling	(3c) Enjoy trolling	(1b) Likely to troll	(2b) Likely to troll	(3b) Likely to troll
Religion (ref. nonreligious)						
Other religion		−0.023	−0.014		−0.052*	−0.051*
		(0.067)	(0.066)		(0.052)	(0.051)
Percentage political posts		0.108**	0.104*		0.229***	0.174***
		(0.001)	(0.001)		(0.001)	(0.001)
Empathy			−0.101*			−0.056
			(0.024)			(0.018)
Big5 Conscientiousness			−0.115**			−0.142***
			(0.002)			(0.002)
Big5 Agreeableness			−0.070			−0.106***
			(0.003)			(0.002)
Big5 Extroversion			0.098**			0.032
			(0.002)			(0.002)
Big5 Neuroticism			−0.061			−0.067*
			(0.002)			(0.002)
Big5 Openness			−0.021			−0.040
			(0.002)			(0.001)
Heavy social media use			0.042			0.010
			(0.034)			(0.024)
Upward online comparison			−0.006			−0.039
			(0.012)			(0.009)
Downward online comparison			0.027			0.156***
			(0.012)			(0.011)
Percentage of similar groups			−0.072*			−0.080***
			(0.000)			(0.000)
Percentage of dissimilar groups			−0.052			0.008
			(0.001)			(0.000)
Male (ref. female or diverse)			0.028			0.019
			(0.025)			(0.020)
Age			−0.066*			−0.030
			(0.001)			(0.001)
LGBTQIA+ (ref. heterosexual)			−0.030			0.003
			(0.031)			(0.023)
Race (ref. white)						
Black			0.009			0.002
			(0.047)			(0.032)
Other			−0.057			−0.015
			(0.034)			(0.027)
Constant	0.187***	0.487***	1.080***	−0.019	0.215***	0.816***
	(0.026)	(0.094)	(0.169)	(0.020)	(0.065)	(0.126)
Observations	909	909	909	909	909	909
R^2	0.32	0.37	0.43	0.51	0.59	0.65

Notes: Standardized beta coefficients, standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

As for identity and political variables, being more engaged in online politics is positively related to both enjoyment of trolling and the likelihood to troll, but being a Republican only increases the likelihood to troll, whereas being a Hindu only makes it more enjoyable. It seems that general political involvement does not differentiate between different types of trolling, but specific political affiliation does. It is not related to the enjoyment of trolling, since if you troll because you defend your position, fun might also not be the main motivation. The significance of Hinduism is related to Hindutva, a form of Hindu nationalism that strongly clashes with the Muslim community (Udupa, 2018) and therefore probably causes some extreme trolling. Religion, however, is not related to the general likelihood to troll.

As for models with further control variables, they do render most trolling motivations insignificant. The only variable maintaining its significance is feeling Powerful and Happy while trolling. Also, Attacked and Annoyed becomes significant with respect to the enjoyment of trolling in the third model, probably due to the stronger irritability and general preparedness of trolls who enjoy trolling. All in all, this means that the impact of personality factors empathy, conscientiousness, and extroversion on the enjoyment of trolling, and of conscientiousness, agreeableness, and neuroticism on the likelihood to troll do not completely confound the role of social or situational determinants. The influence of the former became even stronger (see political affiliations and religiosity in Model 3b). Being involved in more similar groups online stops one from both forms of trolling, whereas downward online comparison increases only the likelihood to troll, which might suggest that it is more related to defensive/reactionary trolling and to increase one's self-worth through trolling. Except for the association between being of younger age and having more fun while trolling, there are no other significant relations between sociodemographics and trolling. This might be due to the numerous other variables we control for in the regressions.

To summarize, the enjoyment of trolling and the likelihood to troll share some of the same determinants but also present some differences. First, defensive motivations are only associated with likelihood to troll. Second, having employment (i.e., a higher social status) increases trolling for fun, whereas precarious working conditions are only related to the likelihood to troll. Third, with the exception of Hinduism and its association with trolling for fun, political affiliation and religiosity do impact only on likelihood to troll. In this regard, we found that people with strong political identities and religious beliefs engage in trolling, but do not necessarily get any satisfaction from it.

6. Conclusion and Discussion

The objective of this study was not so much to demonstrate the untruth of present theories of trolling describing it as malicious behavior, as to give more place to other forms and causes of this behavior—especially the more defensive/reactive forms of trolling, which appear to be partially driven by socioeconomic and sociopolitical factors. The idea of trolls as cold psychopaths who gain pleasure from hurting others is strongly psychology-driven and accounts only for personality traits. Drawing on more general theories of social deviance and negative politics framework, we extended the scope of analysis and found that trolling behavior might also be motivated by social isolation, low social status, and strong political partisanship.

In the pages above, we were able to take some first steps towards analyzing how these socioeconomic and sociopolitical factors impact on two forms of trolling: trolling as entertainment and defensive/reactive trolling. First, we showed that enjoyment of trolling is distinguishable from general trolling behavior and that the latter is partially driven by other factors, such as religiosity or political affiliation. Second, we demonstrated that trolling can be perceived not only as an enjoyable activity motivated by silliness, boredom, and/or a desire to show one's power; it also sometimes occurs in response to an attack by a provoking third party, or due to anxiety or stress, and may leave the trolling individual with regrets. Third, these different trolling motivations can be grouped into two different categories (trolling for fun and defensive/reactive trolling), each of which shows different patterns of correlation (thus, defensive trolling motivations appear to have no connection to the enjoyment of trolling).

Nevertheless, our study has several limitations. Although we were able to show some different underlying motivations for trolling behavior, the main dependent variables used in our analyses still differentiate too poorly between trolling for fun and defensive/reactive trolling. We were only able to distinguish enjoyment of trolling from general trolling behavior—not from defensive trolling in particular. This might also help explain why our regression results only hint at how these different ways of trolling present different patterns of association with respect to social, political, personality, and situational factors, without offering unequivocal evidence for different explanatory patterns. The second reason why some socio-structural variables underperformed in our study is that the survey used for this study was geared towards psychology rather than the social sciences. Hence, some of the crucial socioeconomic variables such as education, class, and income—all of which are standard in sociological or political science research—are missing. This means that we would profit from further analyses with better data in order to solidify these currently preliminary results.

The trolling research would also benefit from a similar study that controlled not for Big5 personality traits, but for dark ones. This would yield far less ambivalent results with respect to trolling behavior. Further research could also attempt to shed more light on how the Attacked and Annoyed motivation relates to different

sorts of trolling, and disentangle the complex and counterintuitive findings of this study. It merits noting that our initial hypothesis—that anger and the experience of being attacked have a stronger association with defensive/reactive trolling than with trolling for fun—turns out not to be the case. Rather, they were negatively related to likelihood to troll (even after controlling for other motivations). By contrast, this factor becomes positively associated with the enjoyment of trolling when analyzed in regression models with more controls. Although we have provided a few possible explanations of these results, a more thorough inquiry would be very useful.

The last limitation we want to mention is how we measured motivation. Our survey asks how one "feels/is" while trolling, not exactly why one is trolling. While the omission of the direct question could be seen as counterproductive, we believe that this slightly concealing formulation is merited due to the risk of rationalization or dishonesty on the part of the respondent if asked directly. Also, the survey makes it nearly impossible to disentangle the causal relationship between the reported emotional state and the action. Did the trolling person feel a certain way (for example, powerful) before starting the trolling behavior, or rather start to feel this way unintentionally while trolling, or begin to troll with the objective to feel powerful? We need research that seeks to separate out these processes, especially since part of our argument distinguishes between trolling as a way of making oneself feel less powerless (defensive/reactive trolling) and trolling for fun from a place of power.

Acknowledgments

We would like to thank Mohammad Al Abdullah, Luzie Vogt, and Widya Puspitasari for the help in preparing this manuscript and organizers as well as participants of the workshop "Negative Politics: Leader Personality, Negative Campaigning, and the Oppositional Dynamics of Contemporary Politics" (November 11–12, 2021) for the comments which tremendously improved the manuscript. We also gratefully acknowledge funding from the Volkswagen Foundation project "Value conflicts in a differentiated Europe: The impact of digital media on value polarization (ValCon)."

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Abts, K., & Baute, S. (2021). Social resentment, blame attribution and Euroscepticism: The role of status insecurity, relative deprivation and powerlessness. *Innovation: The European Journal of Social Science Research*, *35*(1), 39–64.

Adamczyk, A. (2012). Understanding delinquency with friendship group religious context: Delinquency with friendship group religious context. *Social Science Quarterly*, *93*(2), 482–505.

Agnew, R. (1995). Testing the leading crime theories: An alternative strategy focusing on motivational processes. *Journal of Research in Crime and Delinquency*, *32*(4), 363–398.

Andersen, I. V. (2021). Hostility online: Flaming, trolling, and the public debate. *First Monday*, *26*(3). https://doi.org/10.5210/fm.v26i3.11547

Aseltine, R. H., Gore, S., & Gordon, J. (2000). Life stress, anger and anxiety, and delinquency: An empirical test of general strain theory. *Journal of Health and Social Behavior*, *41*(3), 256–275.

Bail, C. A. (2021). *Breaking the social media prism: How to make our platforms less polarizing*. Princeton University Press.

Bonnette, A., Robinson, A., Dailey, S., Ceballos, N. A., & Howard, K. (2019). *Upward social comparisons and posting under the influence: Investigating social media behaviors of U.S. adults with generalized anxiety disorder*. Spotlight on Research. https://www.spotlightonresearch.com/mental-health-research/sormhbonnette19

Bor, A., & Petersen, M. B. (2021). The psychology of online political hostility: A comprehensive, cross-national test of the mismatch hypothesis. *American Political Science Review*, *116*(1), 1–18.

Buckels, E. E., Trapnell, P. D., Andjelovic, T., & Paulhus, D. L. (2019). Internet trolling and everyday sadism: Parallel effects on pain perception and moral judgment. *Journal of Personality*, *87*(2), 328–340.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, *67*, 97–102.

Bulut, E., & Yörük, E. (2017). Digital populism: Trolls and political polarization of Twitter in Turkey ERG. *International Journal of Communication*, *11*, 4093–411.

Campbell, W. K., & Miller, J. D. (Eds.). (2011). *The handbook of narcissism and narcissistic personality disorder: Theoretical approaches, empirical findings, and treatments*. Wiley.

Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone can become a troll: Causes of trolling behavior in online discussions. In C. P. Lee & S. Poltrock (Eds.), *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1217–1230). Association for Computing Machinery.

Cioban, S., Lazăr, A. R., Bacter, C., & Hatos, A. (2021). Adolescent deviance and cyber-deviance: A systematic literature review. *Frontiers in Psychology*, *12*, Article 4416.

Cook, C., Schaafsma, J., & Antheunis, M. (2018). Under

the bridge: An in-depth examination of online trolling in the gaming context. *New Media & Society*, *20*(9), 3323–3340.

Craker, N., & March, E. (2016). The dark side of Facebook®: The dark tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences*, *102*, 79–84.

Festl, R., & Quandt, T. (2013). Social relations and cyberbullying: The influence of individual and structural attributes on victimization and perpetration via the internet. *Human Communication Research*, *39*(1), 101–126.

Fichman, P., & Sanfilippo, M. R. (2015). The bad boys and girls of cyberspace: How gender and context impact perception of and reaction to trolling. *Social Science Computer Review*, *33*(2), 163–180.

Fichman, P., & Sanfilippo, M. R. (2016). *Online trolling and its perpetrators: Under the cyberbridge*. Rowman & Littlefield.

Flores-Saviaga, C., Keegan, B., & Savage, S. (2018). Mobilizing the trump train: Understanding collective action in a political trolling community. In J. Hancock (Ed.), *Proceedings of the Twelfth International AAAI Conference on Web and Social Media* (pp. 82–91). AAAI Press.

Foster, J. D., & Trimm, R. F. (2008). On being eager and uninhibited: Narcissism and approach–avoidance motivation. *Personality and Social Psychology Bulletin*, *34*(7), 1004–1017.

Frischlich, L., Schatto-Eckrodt, T., Boberg, S., & Wintterlin, F. (2021). Roots of incivility: How personality, media use, and online experiences shape uncivil participation. *Media and Communication*, *9*(1), 195–208.

Gidron, N., Adams, J., & Horne, W. (2019). Toward a comparative research agenda on affective polarization in mass publics. *APSA Comparative Politics Newsletter*, *29*, 30–36.

Gidron, N., & Hall, P. A. (2017). The politics of social status: Economic and cultural roots of the populist right. *The British Journal of Sociology*, *68*(S1), 57–84.

Gylfason, H. F., Sveinsdottir, A. H., Vésteinsdóttir, V., & Sigurvinsdottir, R. (2021). Haters gonna hate, trolls gonna troll: The personality profile of a Facebook troll. *International Journal of Environmental Research and Public Health*, *18*(11), Article 5722.

Hagan, J., & Foster, H. (2001). Youth violence and the end of adolescence. *American Sociological Review*, *66*(6), 874–899.

Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research. Language, Behaviour, Culture*, *6*(2), 215–242.

Hare, R. D. (2006). Psychopathy: A clinical and forensic overview. *Psychiatric Clinics of North America*, *29*(3), 709–724.

Harteveld, E., Mendoza, P., & Rooduijn, M. (2021). Affective polarization and the populist radical right: Creating the hating? *Government and Opposition*, *57*(4), 703–727. https://doi.org/10.1017/gov.2021.31

Herring, S., Job-Sluder, K., Scheckler, R., & Barab, S. (2002). Searching for safety online: Managing "trolling" in a feminist forum. *The Information Society*, *18*(5), 371–384.

Hochschild, A. R. (2016). *Strangers in their own land: Anger and mourning on the American right*. New Press.

Hopkinson, C. (2013). Trolling in online discussions: From provocation to community-building. *Brno Studies in English*, *39*(1), 5–25.

Howard, K., Zolnierek, K. H., Critz, K., Dailey, S., & Ceballos, N. (2019). An examination of psychosocial factors associated with malicious online trolling behaviors. *Personality and Individual Differences*, *149*, 309–314.

Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual Differences*, *40*(2), 331–339.

Jones, D. N., & Paulhus, D. L. (2011). The role of impulsivity in the dark triad of personality. *Personality and Individual Differences*, *51*(5), 679–682.

Jones, S., Miller, J., & Lynam, D. (2011). Personality, antisocial behavior, and aggression: A meta-analytic review. *Journal of Criminal Justice*, *39*(4), 329–337.

Juvonen, J., & Graham, S. (2014). Bullying in schools: The power of bullies and the plight of victims. *Annual Review of Psychology*, *65*(1), 159–185.

Lee, B. H. (2018). Explaining cyber deviance among school-aged youth. *Child Indicators Research*, *11*(2), 563–584.

Li, Y.-J., Cheung, C. M. K., Shen, X.-L., & Lee, M. K. O. (2022). When socialization goes wrong: Understanding the we-intention to participate in collective trolling in virtual communities. *Journal of the Association for Information Systems*, *23*(3), 678–706. https://doi.org/10.17705/1jais.00737

Liu, P., Guberman, J., Hemphill, L., & Culotta, A. (2018). Forecasting the presence and intensity of hostility on Instagram using linguistic and social features. In J. Hancock (Ed.), *Proceedings of the Twelfth International AAAI Conference on Web and Social Media* (pp. 181–190). AAAI Press.

Lopes, B., & Yu, H. (2017). Who do you troll and why: An investigation into the relationship between the dark triad personalities and online trolling behaviours towards popular and less popular Facebook profiles. *Computers in Human Behavior*, *77*, 69–76.

Maltby, J., Day, L., Hatcher, R. M., Tazzyman, S., Flowe, H. D., Palmer, E. J., Frosch, C. A., O'Reilly, M., Jones, C., Buckley, C., Knieps, M., & Cutts, K. (2016). Implicit theories of online trolling: Evidence that attention-seeking conceptions are associated with increased psychological resilience. *British Journal of Psychology*, *107*(3), 448–466.

Marchal, N. (2021). "Be nice or leave me alone": An intergroup perspective on affective polarization in online political discussions. *Communication Research*, *49*(3), 376–398.

Matthews, J., & Goerzen, M. (2019). Black hat trolling, white hat trolling, and hacking the attention landscape. In L. Liu & R. White (Eds.), *Companion proceedings of the 2019 World Wide Web Conference* (pp. 523–528). Association for Computing Machinery.

May, D., Bossler, A., & Holt, T. (2012). Predicting online harassment victimization among a juvenile Population. *Youth & Society*, *44*(4), 500–523.

Nai, A. (2021). Fear and loathing in populist campaigns? Comparing the communication style of populists and non-populists in elections worldwide. *Journal of Political Marketing*, *20*(2), 219–250.

Nai, A., & Maier, J. (2018). Perceived personality and campaign style of Hillary Clinton and Donald Trump. *Personality and Individual Differences*, *121*(8), 80–83.

Nai, A., Maier, J., & Vranić, J. (2021). Personality goes a long way (for some): An experimental investigation into candidate personality traits, voters' profile, and perceived likeability. *Frontiers in Political Science*, *3*, Article 636745.

Nevin, A. (2015). *Cyber-psychopathy: Examining the relationship between dark e-personality and online misconduct* [Master's thesis, The University of Western Ontario]. Scholarship@Western. https://ir.lib.uwo.ca/cgi/viewcontent.cgi?referer=&httpsredir=1&article=4531&context=etd

Nyhuis, D., Song, H., & Boomgaarden, H. (2020). Disproportionality in media representations of campaign negativity. *Political Science Research and Methods*, *9*(3), 519–531.

Paasch-Colberg, S., Strippel, C., Trebbe, J., & Emmer, M. (2021). From insult to hate speech: Mapping offensive language in German user comments on immigration. *Media and Communication*, *9*(1), 171–180.

Pao, E. (2015, July 16). Former Reddit CEO Ellen Pao: The trolls are winning the battle for the Internet. *Washington Post*. https://www.washingtonpost.com/opinions/we-cannot-let-the-internet-trolls-win/2015/07/16/91b1a2d2-2b17-11e5-bd33-395c05608059_story.html

Pew Research Center. (2016). *Research in the crowdsourcing age, a case study: How scholars, companies and workers are using Mechanical Turk, a "gig economy" platform, for tasks computers can't handle*. https://www.pewresearch.org/internet/2016/07/11/research-in-the-crowdsourcing-age-a-case-study

Quandt, T., Klapproth, J., & Frischlich, L. (2022). Dark social media participation and well-being. *Current Opinion in Psychology*, *45*, Article 101284. https://doi.org/10.1016/j.copsyc.2021.11.004

Rafalow, M. H. (2015). n00bs, trolls, and idols: Boundary-making among digital youth. In S. L. Blair, P. N. Claster, & S. M. Claster (Eds.), *Sociological studies of children and youth* (Vol. 19, pp. 243–266). Emerald Group.

Reiljan, A. (2020). "Fear and loathing across party lines" (also) in Europe: Affective polarisation in European party systems. *European Journal of Political Research*, *59*(2), 376–396.

Salmela, M., & Capelos, T. (2021). Ressentiment: A complex emotion or an emotional mechanism of psychic defences? *Politics and Governance*, *9*(3), 191–203.

Salmela, M., & von Scheve, C. (2018). Emotional dynamics of right- and left-wing political populism. *Humanity & Society*, *42*(4), 434–454.

Sanfilippo, M. R., Yang, S., & Fichman, P. (2017). *Managing online trolling: From deviant to social and political trolls*. In T. X. Bui & R. Sprague, Jr. (Eds.), *Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 1802–1811). HICSS.

Seidman, G. (2013). Self-presentation and belonging on Facebook: How personality influences social media use and motivations. *Personality and Individual Differences*, *54*(3), 402–407.

Sentse, M., Scholte, R., Salmivalli, C., & Voeten, M. (2007). Person–group dissimilarity in Involvement in bullying and its relation with social status. *Journal of Abnormal Child Psychology*, *35*(6), 1009–1019.

Sest, N., & March, E. (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences*, *119*, 69–72.

Shachaf, P., & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science*, *36*(3), 357–370.

Spears, R. (2017). Social identity model of deindividuation effects. In P. Rössler (Ed.), *The international encyclopedia of media effects*. https://doi.org/10.1002/9781118783764.wbieme0091

Toro, R., García-García, J., & Zaldívar-Basurto, F. (2020). Antisocial disorders in adolescence and youth, according to structural, emotional, and cognitive transdiagnostic variables: A systematic review. *International Journal of Environmental Research and Public Health*, *17*(9), Article 3036.

Udupa, S. (2018). Enterprise Hindutva and social media in urban India. *Contemporary South Asia*, *26*(4), 453–467.

About the Authors

Monika Verbalyte holds a BA in political science from the University of Vilnius (2007) and an MA in European sociology from Freie Universität Berlin (2010). Since September 2020, Monika is a research fellow at Europa-Universität Flensburg in the project "Value Conflicts in a Differentiated Europe" (ValCon). Monika's research interests are emotion sociology, political communication, political attitudes and behavior, European sociology, political scandals, values and norms, affective and political polarization, social network analysis, comparative survey research, discourse methodology, and qualitative interviews.

Christoph Keitel holds a master's degree in European Studies. During his graduate studies, he worked as a student assistant in the research project "Value Conflicts in a Differentiated Europe: The Impact of Digital Media on Value Polarisation in Europe." Following a tutoring position, Christoph is currently on a break from academia, working as a junior project manager for online market research at the Hamburg-based Trend Research GmbH. Christoph's research interests include media studies, political polarization, and behavioural science.

Krista Howard earned her PhD in health psychology and neuroscience in 2010 from the University of Texas at Arlington. She is a tenured professor in the Psychology Department at Texas State University. Dr Howard studies the relationship between mental health and health behavior outcomes, with a focus on social media behaviors, occupational health, chronic illnesses, and maternal health.

POLITICS AND GOVERNANCE
ISSN: 2183-2463

Politics and Governance is an internationally peer-reviewed open access journal that publishes significant and cutting-edge research drawn from all areas of political science.

Its central aim is thereby to enhance the broad scholarly understanding of the range of contemporary political and governing processes, and impact upon of states, political entities, international organisations, communities, societies and individuals, at international, regional, national and local levels.

cogitatio

www.cogitatiopress.com/politicsandgovernance